

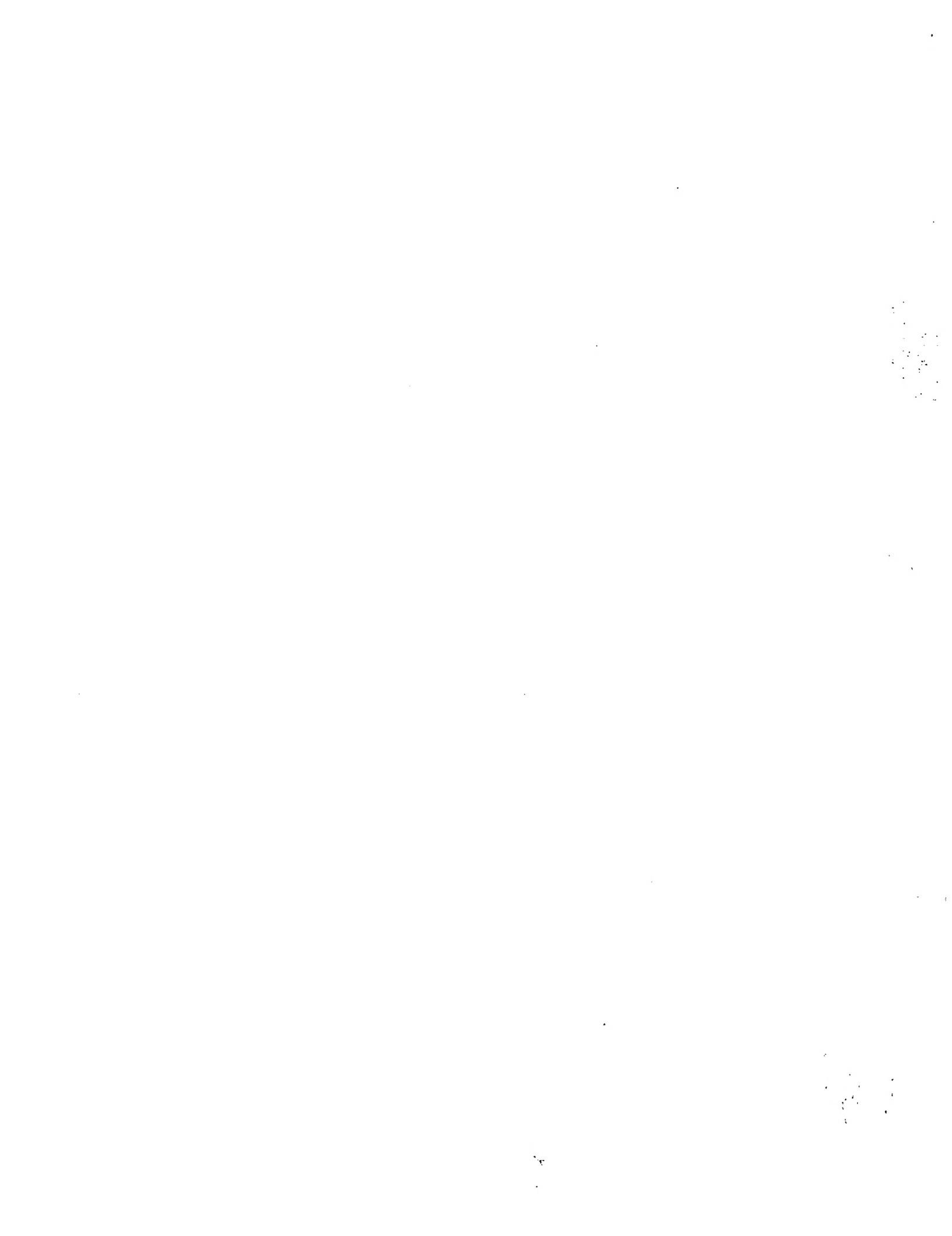
**begin set** ↪ FINDPATTERNS on  $\text{pir}[*]$  allowing 0 mismatches

**DRX>** end of seq; x = any A  
Databases searched: NBRF, Release 73.0, Released on 16Aug2002, Formatted on 20Aug2002  
Total finds: 0  
Total length: 95,134,422  
Total sequences: 283,224  
CPU time: 01:02.36

The Quest program provides the answer sets for the issued, pending & A-Genesis files. The reference & alignment data are mapped to sections in these files.

Finally, there are 2 disks for seg 1 run on computer w/ the dx scoring table.

Ex. Patterson, find  
The GC/G program, find patterns,  
(it) searches both the PIR & Swissprot  
data bases. At first, I searched  
the sequences w/ no limitations on  
length of the database sequences.  
The files were huge & unmanage-  
able. Then I searched exact  
length (see left). There were ~~no~~  
answers in SWISSPROT or PIR for  
the exact length seq.  
Then I searched the sequence  
w/ 20 AA's on either side. These  
answers were sets for each database  
base. The seq 2 spl. find (a  
seq 2 pir1. find) file shows the  
alignments & the accession numbers  
of the database are seq. The .ref  
files give the information on the  
accession number. (Sorry I can't  
many them together.)





1 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{8}{T}DRx{6})NFLGPFCRDKNSFRI

1: Q9KIL6 ck: 2042 len: 24 ! Q9k1l6 streptomyces coelicolor a3(2). f81r  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{A}{A}DRx{19})MADRGTDDEATAERALGSRapef

1: Q9x9u4 ck: 4039 len: 25 ! Q9x9u4 streptomyces coelicolor. adp glucose  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{18}{A}DRx{4})MEGIVLAGGESKRMLPMTABRKA

1: Q49137 ck: 9647 len: 32 ! Q49137 methylobacterium extorquens. myas pr  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{16}{A}DRx{13})AURICAPFARPFPLRADRFDAEALSRHLMTT

1: Q53914 ck: 3396 len: 9 ! Q53914 streptomyces cyanus (streptomyces c  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{1}{T}DRx{5})ITDRAAQOP

1: Q50476 ck: 7152 len: 13 ! Q50476 mycobacterium tuberculosis. catalase  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{8}{T}DRx{2})MPEQHPPIITDRSR

1: Q05422 ck: 6757 len: 38 ! 005422 mycobacterium marinum. phytoene dehy  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{15}{A}DRx{19})VFGVGVPFTLISGRLLAADRITGNNTKSIRHDLKAQLS

1: Q93A12 ck: 9419 len: 22 ! Q93a12 thiobacillus ferrooxidans. putative  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{9}{D}DRx{10})AONFINPDLDDETEKDLATSTV

1: Q93a08 ck: 5753 len: 12 ! Q93a08 thiobacillus ferrooxidans. resb prot  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{2}{Q}{D}DRx{6})OSODDRKENNDG

1: Q03977 ck: 4923 len: 19 ! Q03977 escherichia coli. alpha-acceptor pol  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{13}{T}DRx{3})MMITDSIΛAVARTDRPSQ

1: Q99094 ck: 1642 len: 36 ! Q99094 salmonella typhimurium. mkaa protein  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
 x{15}{D}DRx{18})

1: Q9N225 ck: 6136 len: 20 ! Q9n225 homo sapiens (human). atp7b (frag  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{6}{A}DRx{11})APIQQLADRFSGYFVFPIVI

1: Q15244 ck: 2868 len: 40 ! Q15244 homo sapiens (human). phosphoribo  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{19}{A}DRx{18})MPNIKIFSGSSHSQDLSQKIA DRIGLEKGKVWTKFSNQET

1: Q9BQV8 ck: 5960 len: 34 ! Q9bqv8 homo sapiens (human). phosphoribo  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{19}{A}DRx{18})NGIITKTDKVFEVMLATDORSHYAKCNPYMDSPSI

1: Q9UCI1 ck: 8961 len: 15 ! Q9uci1 homo sapiens (human). tropomyosin  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{2}{E}{A}DRx{7})HIAEDADRKYEVAR

1: Q96F68 ck: 588 len: 28 ! Q96f68 homo sapiens (human). unknown (pr  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{10}{A}DRx{15})VSQPGSCRIGADRVGHVGQORAGGVRE

1: Q9NBB1 ck: 6662 len: 26 ! Q9nbbl drosophila melanogaster (fruit fl  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{8}{T}DRx{15})VVARNFFGTDRIFVTVTKIPKKKEE

1: Q9GU45 ck: 9325 len: 27 ! Q9gu45 sycon raphanus. homeobox protein  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{20}{T}DRx{4})KTSKYLA PTFRQALARLGLTDROQVK

1: Q9GU44 ck: 9332 len: 27 ! Q9gu44 sycon raphanus. homeobox protein  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{20}{T}DRx{4})RTSKYLA PTFRQALARLGLTDROQVK

1: Q9xzz6 ck: 1592 len: 23 ! Q9xzz6 littoria saxatilis. calmodulin (br  
 <X{0..20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
 x{8}{A}DRx{11})LDMINEVADQRDDRLPRPVY

```

Q94681 ck: 5282 len: 25 ! Q94681 polyandrocarpa misakiensis. pmhbbox1
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[20]{(T)DRX{[2]} }}}}}} }}>
1: HFNOQLCERRRQEVAKAVNLTDRQV

1:
Q23749 ck: 8475 len: 27 ! Q23749 ctenodrilus serratus. ovxl ortholog
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[19]{(T)DRX{[5]} }}}}}} }}>
1: POQEILVTDGTIAHASPETDRGSGEN

1:
Q25482 ck: 9269 len: 27 ! Q25482 metridium senile (brown sea anemone)
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[20]{(T)DRX{[4]} }}}}}} }}>
1: HFNFHLIKERSEMATORLNLTDRQVKI

1:
Q94458 ck: 8578 len: 27 ! Q94458 chaetopterus variopedatus. chv-hb11
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[20]{(T)DRX{[4]} }}}}}} }}>
1: XNOKFIERKDRDRISNETGLDDROKI

1:
096892 ck: 9212 len: 27 ! 096892 sacculina carcin. caudal (fragment)
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[20]{(T)DRX{[4]} }}}}}} }}>
1: RENNITYIKRKLELSRLTGLTDQRQVKI

1:
P91901 ck: 9410 len: 27 ! P91901 beroe ovata. homeobox protein bpx26
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[20]{(T)DRX{[4]} }}}}}} }}>
1: LFNMWLTERRLEISKINLNDRQVKI

1:
002602 ck: 9552 len: 27 ! 002602 beroe ovata. homeobox protein bpx35
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F, H, N
x{[20]{(T)DRX{[4]} }}}}}} }}>
1: XNOKFIERKDRDRISNETGLDDROKI

1:
Q9TRF2 ck: 1454 len: 33 ! Q9trf2 bos taurus (bovine). vamp/synapto
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F,
x{[2]{(A, A)DRX{[8]} }}}}}} }}>
1: XEAAADREDDPNFF

1:
Q9TR40 ck: 8661 len: 15 ! Q9tr40 bos taurus (bovine). glutamate de
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F,
x{[2]{(A, A)DRX{[8]} }}}}}} }}>
1: XEAAADREDDPNFF

1:
Q9TRF4 ck: 1454 len: 33 ! Q9trf2 bos taurus (bovine). vamp/synapto
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F,
x{[17]{(D)DRX{[13]} }}}}}} }}>
1: XVNVDKVLERQKLSELDDRADALQAGASQFET

1:
Q37112 ck: 9119 len: 22 ! Q37112 pinus thunbergii (green pine) (ja
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F,
x{[5]{(T)DRX{[14]} }}}}}} }}>
1: MEYLTIDRSIECGIVIRKIEST

1:
Q957T4 ck: 9875 len: 22 ! Q957t4 abies alba (edeltanne) (european
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F,
x{[17]{(T)DRX{[2]} }}}}}} }}>
1: MPTSNOSIRHGREGKRRTDRTR

1:
Q957T0 ck: 1370 len: 23 ! Q957t0 pinus mugo. ribosomal protein s12
<X{[0, 20]{(A, D, E, F, G, I, N, S, V){[0, 1]{(A, D, E, H, I, L, M, Q, S){[0, 1]{(D, E, F,
x{[17]{(T)DRX{[3]} }}}}}} }}>
1: MPTSNOSIRHGREGKRRTDRTR

```

Q95758	ck: 1370 len: 23 ! Q95758 pinus sylvestris (scots pine). ribos	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	MPTSNOSTHGREKKRRRTDTRA	x{11}{T}DRX{3}
Q9t2k6	ck: 5922 len: 20 ! Q9t2k6 cucurbita moschata (cushaw squash)	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{3}{S}{L}{A}{DRX{12}}	x{11}{D}DRX{2}
1:	GSSSLADRLXLGLSLAXDGFS	NNTNTQMTAYDDDRGX
Q9t2h3	ck: 2853 len: 24 ! Q9t2h3 spinacia oleracea (spinach). chaperon	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{14}{A}{DRX{7}}	x{5}{A}DRX{4}
1:	ATWVAPKYSKIPFADRVLIKIE	NLANIKADRODE
Q37852	ck: 3123 len: 18 ! Q37852 bacteriophage r17. a protein (fragment)	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{8}{E}{DRX{7}}	x{6}{T}DRX{8}
1:	ALRYLAJNEDRKFRSKH	MPVKKRDTRBALSLEY
Q42209	ck: 1580 len: 28 ! Q42209 arabidopsis thaliana (mouse-ear cress)	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{8}{T}DRX{17}	x{4}{A}DRX{20}
1:	PEAHMVKTDRAKVESTARSWIQKYAM	MAENIDRNQIEKLNNRKEEVERL
Q9fzp4	ck: 3486 len: 29 ! Q9fpz4 arabidopsis thaliana (mouse-ear cress)	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{20}{E}{DRX{6}}	x{12}{T}DRX{2}
1:	MDVDDDFSYLLSKEIDEENEDREPKVY	LFNMYLTERRLEISKINTLTDROV
Q39633	ck: 1444 len: 23 ! Q39633 cucumis sativus (cucumber). catalase	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{4}{A}{DRX{16}}	x{17}{T}DRX{13}
1:	NESQDADSVGQKLAPHLNVRPST	IHTGEKPYRCAECKAFDRSNLKHQPTHTGE
P82195	ck: 7110 len: 26 ! P82195 spinacia oleracea (spinach). chlorophyll a	1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{6}{E}{DRX{17}}	x{10}{T}DRX{18}
1:	KAHTRREDRTARHVRTRKKVSGTPER	ALQAFEFEHGATDRGSPALSSQALVRRVLDN
P82196	ck: 4154 len: 29 ! P82196 spinacia oleracea (spinach). chlorophyll a	1 Q9q46 ck: 4770 len: 25 ! Q9q46 mus musculus (mouse). amyloid precursor protein
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{6}{E}{DRX{20}}	x{11}{A}{DRX{10}}
1:	KAHTRREDRTARHVRTRKKVSGTPER	ENEVEPVDAKPAAADRGLTTRPGSGL
P83089	ck: 4307 len: 19 ! P83089 spinacia oleracea (spinach). thylakocyanin	1 Q9et00 ck: 5807 len: 20 ! Q9et00 mus musculus (mouse). eif4h (frag)
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{12}{T}DRX{4}	x{7}{A}{DRX{10}}
1:	ATANAPLILITITDRVFFD	MADDITYDORAYSFGGGGK
Q9s8d2	ck: 2605 len: 29 ! Q9s8d2 cucumis melo (muskmelon). cmeti-b-transcript	1 Q9z115 ck: 4927 len: 19 ! Q9z115 rattus rattus sp. calcineurin a alpha (e
1:	<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,	
1:	x{13}{T}DRX{3}	MSEPKATIDPKLSTTDRVVK

Q62256 ck: 3193 len: 18 ! Q62256 mus musculus (mouse). spermatogenic-  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..2}{D}DRX{5})  
1: MSGKQDSFWEDRIPGR

Q63985 ck: 5721 len: 25 ! Q63985 rattus sp. hox1.8 homeobox homolog F  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{2..21}{T}DRX{2})  
1: LFNNYLTRERRELEISRSVHLDROV

Q61159 ck: 837 len: 32 ! Q61159 mus musculus (mouse). nn8-ag (fragme-  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..15}{A}DRX{14})  
1: MAASMCDFVFSCFGVADRGARSVEVYVDSIK

Q923H1 ck: 2032 len: 40 ! Q923h1 cricetus cricetus (chinese hamster)  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..19}{A}DRX{15})  
1: PIPAVEGAVASGALADPATAADRASSIAFXGSRPRSMW

Q99KX5 ck: 9106 len: 27 ! Q99KX5 mus musculus (mouse). hypothetical 2  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..12}{A}DRX{12})  
1: MASLPPVVGQSYPADRGFLGLPRAVOEH

Q9QVD0 ck: 3385 len: 24 ! Q9qvd0 cavia (guinea pigs). vitronectin (fibrinogen)  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..14}{A}DRX{7})  
1: XESXKGRXTCFNADRXXQXEL

Q9QV89 ck: 581 len: 32 ! Q9qv89 rattus sp. slp-14\*fatty acid-binding  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..14}{D}DRX{14})  
1: EISFOLGVETDEVTAADRKYKSVVILDGGLKLV

Q9QV71 ck: 1363 len: 17 ! Q9qv71 rattus sp. lactase-phlorizin hydrolase  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
(D,E)DRX{12})  
1: DWEDRNFTIAGPLTNDL

Q61461 ck: 4925 len: 25 ! Q61461 mus musculus (mouse). cytochrome P-450 15A1  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..17}{D}DRX{15})  
1: SHRLPLTDDRSRKMPTDAVHIEOR

Q65747 ck: 4484 len: 19 ! Q65747 bluetongue virus. outer coat protein  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,H,N  
x{1..13}{E}DRX{2})  
1: ERLKIFEHRRNRDDERFY

Q66538 ck: 8137 len: 35 ! Q66538 ebola virus. 3' proximal protein (fragment)  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..6}{N}DRX{3})

1: <X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..2}{D}DRX{20})  
1: MRKINNFSLSKFDDRNLLKLLICNHVTWSEPHTS

Q83622 ck: 3517 len: 9 ! Q83622 murray valley encephalitis virus.  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..4}{E}DRX{2})  
1: THVSEDML

Q84254 ck: 9056 len: 22 ! Q84254 bovine papillomavirus. x protein  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..3}{A}DRX{6})  
1: NAGPKPGHTPPEDVADRPPLPE

Q9PWCl ck: 9645 len: 39 ! Q9pwcl brachydanio rerio (zebrafish)  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..19}{D}DRX{17})  
1: LYELLSPVLSPLOPHVYESPDDRSFLHAMFGERSLHSLVR

093438 ck: 3493 len: 29 ! 093438 gallus gallus (chicken). lbxl1 pro-  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..10}{A}DRX{16})  
1: RFLQYKYLSPADRDQIAQOLQLGLTNQAVIT

Q90297 ck: 9410 len: 27 ! Q90297 carassius auratus (goldfish). hom-  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..20}{T}DRX{4})  
1: LFNNYLTRERRELEISRSVHLDROVKT

Q90298 ck: 9642 len: 27 ! Q90298 carassius auratus (goldfish). hom-  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..20}{T}DRX{4})  
1: LFNNYLTRERRELEISRSVHLDROVKT

Q07145 ck: 9552 len: 27 ! Q07145 petromyzon marinus (sea lamprey).  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..20}{T}DRX{4})  
1: LFNNYLTRERRELEISRSVHLDROVKT

Q07150 ck: 9407 len: 27 ! Q07150 petromyzon marinus (sea lamprey).  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..20}{T}DRX{4})  
1: LFNSMYLURERRELEISRSVHLDROVKT

073591 ck: 8165 len: 14 ! 073591 gallus gallus (chicken). hox.c10  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..10}{T}DRX{4})  
1: REIJSKSNLNDRO

P83009 ck: 7125 len: 13 ! P83009 gallus gallus (chicken). phospholeman (fxva)  
<X{0..20}(A,D,E,F,G,I,N,S,V){0..1}(A,D,E,H,I,L,M,Q,S){0..1}(D,E,F,  
x{1..6}{N}DRX{3})

Sun Dec 22 19:28:25 2002

seq2swpl.find

1: VSDVPNNDDRFY  
P83010 ck: 6920 len: 13 ! P83010 triakis scyllium (leopard shark) (tr  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{17}{S}DRX{3}  
1: AGEPANNEDRENY  
Q90ZG3 ck: 9516 len: 27 ! Q90zg3 oryzias latipes (medaka fish). hexch  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{20}{T}DRX{4}  
1: FENVYINKEKRLQLSRMLNLTDRQVKI  
Q9PS38 ck: 6634 len: 20 ! Q9ps38 rana catesbeiana (bull frog). po g1x  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{4}{T}DRX{13}  
1: IEVYTDRQEISNVGSKVHLY  
Q9PS32 ck: 1521 len: 23 ! Q9ps32 gallus gallus (chicken). nonmuscle a  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{16}{E}DRX{4}  
1: LASDLEWIRRTPWLEDRSPOK  
Q9PRV3 ck: 2489 len: 24 ! Q9prv3 gallus gallus (chicken). nap-185 neu  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{2}{G}{E}DRX{18}  
1: DEGEDRDAVERAVLGNGGCRRPK  
Q9DU42 ck: 5339 len: 25 ! Q9du42 human immunodeficiency virus type 1.  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{5}{E}DRX{17}  
1: MAGRSSEDRDEELKTVRLKLYQS  
Q70140 ck: 3335 len: 9 ! Q70140 human immunodeficiency virus type 1.  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{4}{T}DRX{2}  
1: SKRETBDRED  
Q25575 ck: 575 len: 23 ! Q25575 helicobacter pylori (campylobacter f  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{13}{E}DRX{1}  
1: MSPLTPLRNPLTQEDRFQEQITA  
Q06283 ck: 8577 len: 27 ! Q06283 mycobacterium tuberculosis. hypothet  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{11}{A}DRX{12}  
1: MPASSIGTGSPADRIDATHERRREV  
Q9KTZ5 ck: 3787 len: 33 ! Q9ktz5 vibrio cholerae. hypothetical protein  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{15}{A}DRX{15}  
1: MPVGSSTPCSTALADRKYVVSRYWCFFTLVIL

Databases searched:  
SWISS-PROT, Release 40.3, Released on 9Aug2002, Formatted on 20Aug2002  
SPREMBL, Release 21.0, Released on 15Jun2002, Formatted on 28Jun2002  
Total finds: 118  
Total length: 247,523,443  
Total sequences: 784,472  
CPU time: 04:50:37

Q9KL13 ck: 1300 len: 32 ! Q9kl13 vibrio cholerae.  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{13}{T}DRX{16}  
1: MPDLURFSACSTDRRKQTNLFSSDLKQQWR  
Q8X407 ck: 6940 len: 13 ! Q8x407 escherichia coli o157:h7. hypothesis  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{S}{T}DRX{8}  
1: MSTDRRKPVMLIFH  
Q8VJ09 ck: 8356 len: 35 ! Q8vj09 mycobacterium tuberculosis. hypothesis  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{13}{A}DRX{19}  
1: MYKLAIDVEYDADRTERGAILRRENLYSSLTE  
Q9HNL3 ck: 9872 len: 32 ! Q9hn13 halobacterium sp. (strain nrc-1).  
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{17}{A}DRX{12}  
1: MVRVPVTDAGRAEMERADRLETTAAFRNLVD

```

! FINDPATTERNS on pir.* allowing 0 mismatches
! 1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N,M,P,V,W}{0
A21630 ck: 4925 len: 25 ! cytochrome P450b - mouse (fragment)
1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{17}{D}DRX{15}}
1: SHRLPILDDRSKMPYDAVHEIQR

FA2762 ck: 5572 len: 20 ! C 3.4.25.1 proteasome endopeptidase complex
1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{9}{A}DRX{7}}
1: MKGKNXVAIAADRFFGQAQ

S55729 ck: 1645 len: 28 ! orotidine-5'-monophosphate decarboxylase -
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{16}{A}DRX{9}}
1: VTVSDADYTITKELLDADRIVGFVSTRS

B44336 ck: 2946 len: 34 ! neurotoxin Tx3-2 - spider (Phoneutria nigril
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{17}{E}DRX{14}}
1: ACAGLYKKKGKGASPCEDRCKCDLAMGNICK

B28563 ck: 9362 len: 22 ! hemoglobin chain IV - earthworm (Lumbricus
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{9}{E}DRX{10}}
1: ADEBESCCSYEDRREVRHIDDDW

S01808 ck: 3406 len: 24 ! hemoglobin AII - tube worm (Lamellibrachia
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{17}{E}DRX{14}}
1: SNSNCTTEDRQMOLMWANVNSAQ

D60894 ck: 6311 len: 20 ! gamma crystallin V - bullfrog (fragment)
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{6}{E}DRX{11}}
1: GKVIVFEDRNQGRSYECSS

S04621 ck: 5853 len: 20 ! gamma-crystallin II - milk shark (fragment)
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{16}{E}DRX{11}}
1: GRKTFYEDRGFOGHCYECSS

A61570 ck: 6107 len: 20 ! gamma-crystallin (total) - siamang (fragment)
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{16}{E}DRX{11}}
1: GRKTFYEDRGFOGHRYCCTT

B61570 ck: 6058 len: 20 ! gamma-crystallin (total) - Formosa sika deer
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N
x{16}{E}DRX{11}}
1: GKTFYEDRGFOGRHYCCTT

A60502 ck: 4514 len: 25 ! myonexin - northern leopard frog (fragment)
1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{11}{E}DRX{11}}
1: EDGDFPEYDGRVIFSLKNYKA

E33208 ck: 2821 len: 29 ! calreticulin, uterine - rabbit (fragment)
1 <X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{16}{T}DRX{10}}
1: EPVVYFKEQFLDGDGTWIESKHKSDF

C33208 ck: 2871 len: 29 ! calreticulin, slow twitch skeletal muscle
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{16}{T}DRX{10}}
1: EPVVYFKEQFLDGDGTWIESKHKSDF

A33434 ck: 2817 len: 24 ! calcium-binding protein - dog (fragment)
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{15}{T}DRX{5}}
1: EPAIVFKEQLFDGGXTRXIESK

B41481 ck: 1642 len: 36 ! virulence-associated protein (vira 5' re
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{16}{T}DRX{5}}
1: GSQDFASQLKRSLRSDDRTADTNRIKRINMRVLNS

S07394 ck: 9004 len: 22 ! DNA-invertase - phage D108 (fragment)
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{13}{D}DRX{6}}
1: YKKHPAKRHTEENDRINQIDR

S35552 ck: 9008 len: 32 ! vesicle-associated membrane protein 1 -
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{9}{D}DRX{20}}
1: ERDKQLSELDRADALQAGASQFESSAAKLKR

S35553 ck: 9118 len: 32 ! vesicle-associated membrane protein 1 -
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{9}{D}DRX{20}}
1: ERDKQLSELDRADALQAGASQFESSAAKLKR

S35555 ck: 9033 len: 32 ! vesicle-associated membrane protein 2 -
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{9}{D}DRX{20}}
1: ERDKQLSELDRADALQAGASQFESSAAKLKR

S35554 ck: 9033 len: 32 ! vesicle-associated membrane protein 2 -
<X{0,20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,
x{9}{D}DRX{20}}
1: ERDKQLSELDRADALQAGASQFESSAAKLKR

E44621 ck: 9552 len: 27 ! homeotic protein Hox 10 (clone 10w) - se

```

1           <X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
1           x{20}{T}DRx{4}  
1 :           LFFNMYLTERRLEISRGVNVNLIDRQVKI

F44621 ck: 9407 len: 27 ! hypothetical protein Box 10 (clone 10x) - sea 1  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{20}{T}DRx{4}  
1 :           LFSMLTTRRLBISHLISLTDRQVKI

C84355 ck: 9872 len: 32 ! hypothetical protein Vng2049c [imported] -  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{17}{A}DRx{12}  
1 :           MVPYPTDAGREMERADRLTTAAFWRLVD

E64634 ck: 575 len: 23 ! hypothetical protein Hp0917 - Helicobacter  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{13}{E}DRx{7}  
1 :           MSFLTPLRNPLTOEDRFQQEITA

PH0858 ck: 807 len: 23 ! Maud protein - Paracoccus denitrificans (fr  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
(L)(E)(A)DRx{18}  
1 :           LEADRDTGFAASLQOYMASRKQKAA

A61412 ck: 8072 len: 27 ! methane monooxygenase (EC 1.14.13.25) prote  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{12}{A}DRx{12}  
1 :           WIEDYASRIDFKADRDQIVKAVLAGLK

A82288 ck: 3787 len: 33 ! hypothetical protein VC0735 [imported] - V  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{15}{A}DRx{15}  
1 :           MPVGSSIPCSISTALADRKVVSRYWCFTTLVL

BB2421 ck: 1300 len: 32 ! hypothetical protein VCA0761 [imported] - V  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{13}{T}DRx{16}  
1 :           MDPDLRLFSAICSTDRRQQTNLFLSSDLKQQR

B39089 ck: 6231 len: 20 ! hydrogenase (EC 1.18.99.1) 34K chain - Thio  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{4}{A}DRx{13}  
1 :           XXEODADPSVIGLPLFQEXTG

H70954 ck: 8577 len: 27 ! hypothetical protein RV3599c - mycobacteriu  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{11}{A}DRx{12}  
1 :           MPASSLGTGSPAADRIDATHERRREV

S14161 ck: 5957 len: 20 ! probable dtDP-4-keto-6-deoxy-hexose 3,4-is  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,H,N  
x{3}{T}DRx{14}

1 :           MTTDTRAGLGROLOMTRGLH  
S66213 ck: 1672 len: 17 ! glucose 1-dehydrogenase (EC 1.1.1.47) -  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{9}{E}DRx{5}  
1 :           MKIAIAVKRGEDRPVWIE

S23180 ck: 7956 len: 31 ! carboxypeptidase - Sulfolobus solfataric  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{14}{E}DRx{14}  
1 :           MDLVKEKLNEVKELEDRIQIRRKIXENPL

T10123 ck: 1444 len: 22 ! hypothetical protein 22, Psba 5'-region  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{5}{T}DRx{14}  
1 :           MEYLJLTDRIECGJYLLKKTEST

PS0212 ck: 8963 len: 15 ! probable catalase (EC 1.11.1.6) - cucumb  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{4}{A}DRx{16}  
1 :           NGSQADRSVGQKLAAPHLNVRPSI

B60698 ck: 2526 len: 29 ! 29K protein 4228 - rice (strain Nihonbar  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{13}{E}DRx{9}  
1 :           SPADDRDVGDRAD

S19614 ck: 1820 len: 17 ! globin - polychaete (Eudistylia vancouver  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{7}{E}DRx{7}  
1 :           GMKXXSMEDRKTVALDW

A60741 ck: 4352 len: 25 ! insulin-like growth factor-binding prote  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{20}{E}DRx{2}  
1 :           LAFCPGPGQVQAGRPGGVVEEDRGG

B28457 ck: 4360 len: 19 ! proteoglycan II, bone - human (tentative  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{11}{D}DRx{5}  
1 :           DEAXGIAPEVPDDDRPEPS

C54037 ck: 1565 len: 29 ! splicing regulatory protein SWAP homolog  
<X{0..20}{A,D,E,F,G,I,N,S,V}{0,1}{A,D,E,H,I,L,M,Q,S}{0,1}{D,E,F,  
x{17}{E}DRx{9}  
1 :           POLJQEELEAKQAKQKLEDRLAAAREKL

A48845 ck: 9975 len: 23 ! sterol regulatory element 1 binding protein  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{13})(A)DRx{7}  
1: RGRANGIDARFRAGADRGAMDDTF

S22228 ck: 4810 len: 25 ! vitronectin - dog (fragment)  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{14})(A)DRx{8}  
1: LAQESXKGRVTEGFNADRQQDELX

A26393 ck: 7 len: 16 ! annexin 36K chain - pig (fragment)  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{17})(E)DRx{6}  
1: MLLGTTEDDRLLERI

S68260 ck: 6375 len: 38 ! hypothetical protein gadd7.1 - long-tailed  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{18})(E)DRx{17}  
1: MTPSFSGSSKQLQRNAQMEDRGPNNPSEFTTGDNLKT

A35678 ck: 3193 len: 18 ! hypothetical protein (proenkephalin 5' regi  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{10})(E)DRx{5}  
1: MSSGKQDSPFWEDRIPPPR

PH0780 ck: 8887 len: 15 ! T-cell receptor alpha chain (c11) - mouse (c  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{8})(A)DRx{4}  
1: CALSETGGADRIFTG

PH0794 ck: 1531 len: 17 ! T-cell receptor alpha chain (K1 v-alpha 3-E  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{10})(A)DRx{4}  
1: CAVSNMEYRGADRIFTG

PH0806 ck: 8864 len: 15 ! T-cell receptor alpha chain (PE5.1.v-alpha  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{8})(A)DRx{4}  
1: CALSPOGGADRIFTG

S03505 ck: 6013 len: 20 ! T-cell receptor alpha chain J region (80)-  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{4})(A)DRx{13}  
1: NTEGADRITLGKGTLIQP

PT0212 ck: 3993 len: 10 ! T-cell receptor alpha chain V-J region (4-1  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N  
X{6})(A)DRx  
1: CAVAGGADR

PT0547 ck: 2627 len: 8 ! T-cell receptor beta chain V-D-J region (12  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,H,N

1: MSTDRKPYMLFH

1: x(S)(S)(A)DRx  
ASSDADRG

PT0676 ck: 2049 len: 7 ! T-cell receptor beta chain V-D-J region  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{2}(G)(E)DRx  
ASGEDRG

PT0576 ck: 2112 len: 7 ! T-cell receptor beta chain V-D-J region  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X(S)(S)(D)DRx  
ASSDORT

PT0366 ck: 480 len: 28 ! T-cell receptor beta chain V-J region (6  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{7})(E)DRx{18}  
1: LYFCASSEDNRNQLRFLERGLDFSVLED

A46592 ck: 1363 len: 17 ! lactase-phlorizin hydrolase, mitochondrial [v  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
(D)(W)(E)DRx{12}  
DWEDRNFIAGPLTNLD

S78414 ck: 6086 len: 26 ! ribosomal protein RL25, mitochondrial [v  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{19})(E)DRx{4}  
1: ISRRXEKKNKIVYPDQDLGDERRDAE

S22227 ck: 5585 len: 25 ! vitronectin - guinea pig (fragment)  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{14})(A)DRx{8}  
XXESXKXRXTGCFNADRKKQXELX

A54226 ck: 6170 len: 26 ! light-harvesting protein B-830 alpha-1 c  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{18})(D)DRx{5}  
MKVPVMMADENAKLNNEPDDRKKFFV

B54226 ck: 634 len: 23 ! light-harvesting protein B-830 alpha-2 c  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{18})(D)DRx{2}  
1: MOVPVMMADENAKLNNEPDDRKKFFV

T46593 ck: 6757 len: 38 ! phytoene dehydrogenase [imported] - myco  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X{15})(A)DRx{19}  
VPGVPIPTLSGRALADRITGNTTSIRHLDKAQSL

H85575 ck: 6940 len: 13 ! hypothetical protein Z0899 [imported] -  
<X{0,20}(A,D,E,F,G,I,N,S,V){0,1}(A,D,E,H,I,L,M,Q,S){0,1}(D,E,F,  
X(S)(T)DRx{8}  
1: MSTDRKPYMLFH

PA0041 ok: 9117 len: 15 ! plastoquinol-plastocyanin reductase (EC 1.1  
1 <X{0..20}{A,D,E,F,G,I,N,S,V}{0..1}{A,D,E,H,I,L,M,Q,S}{0..1}{D,E,F,H,N  
x{5}{A})DRx{7}  
1: ASSIPADRPDMKRR

Databases searched:

NBRF, Release 73.0, Released on 16Aug2002, Formatted on 20Aug2002

Total finds: 63  
Total length: 96,134,422  
Total sequences: 283,224  
CPU time: 01:39.36

!AA\_SEQUENCE 1.0  
 F1;A21630 - cytochrome P450b - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Accession: B44336  
 C;Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 11-Jun-1999  
 R;Stupans, I.; Ikeda, T.; Kessler, D.J.; Nebert, D.W.  
 A;Title: Characterization of a cDNA clone for mouse phenobarbital-inducible cytochrome P450b.  
 A;Reference number: A21630; MUID:84207435; PMID:6547088  
 A;Accession: A21630  
 A;Molecule type: mRNA  
 A;Residues: 1-25 <STD>  
 A;Cross-references: GB:K02409; NID:9192896; PID:AAA37510\_1; PID:g553905  
 A;Note: the authors translated the codon CTA for residue 5 as Pro, ACC for residue 7 as Ser and AGT for residue 12 as Thr  
 C;Superfamily: human cytochrome P450 CYB2D6; cytochrome P450 homology  
 C;Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protein

A21630 Length: 25 December 22, 2002 19:21 Type: P Check: 4925 ..

1 SRLPTLDDR SKMPTYDAVI HETQR

!AA\_SEQUENCE 1.0  
 F1;F42762 - C 3.4.25.1 proteasome endopeptidase complex () subunit 13 - bovine (fragment)  
 C;Species: Bos primigenius taurus (cattle)  
 C;Accession: F42762  
 R;Dick, L.R.; Moonaw, C.R.; Pramanik, B.C.; DeMartino, G.N.; Slaughter, C.A.  
 A;Title: Identification and localization of a cysteiny1 residue critical for the trypsin-like catalytic activity of the proteasome  
 A;Reference number: A42762; MUID:92378961; PMID:1510924  
 A;Accession: F42762  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-20 <DCI>  
 A;Note: sequence extracted from NCBI backbone (NCBIP:112175)  
 C;Superfamily: multicatalytic endopeptidase complex chain C9  
 C;Keywords: hydrolase

F42762 Length: 20 December 22, 2002 19:21 Type: P Check: 5572 ..

1 MKGKXVAIA ADRRFGIQAQ

!AA\_SEQUENCE 1.0  
 F1;S55729 - orotidine-5'-monophosphate decarboxylase - Aspergillus awamori (fragments)  
 C;Species: Aspergillus awamori  
 C;Accession: S55729  
 C;Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
 C;Accession: S55729  
 R;Gouka, R.J.; Hessing, J.G.M.; Stam, H.; Musters, W.; van den Hondel, C.A.M.J.J.  
 Curr. Genet. 27, 536-540, 1995  
 A;Title: A novel strategy for the isolation of defined pyrg mutants and the development of a site-specific integration system for Aspergillus awamori.  
 A;Reference number: S55729; MUID:9601709; PMID:7533938  
 A;Accession: S55729  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-18-19-28 <GOU>  
 C;Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate decarboxylase homology

S55729 Length: 28 December 22, 2002 19:21 Type: P Check: 1645 ..

1 VTVSADWVIT KELLDLADRV MGFWSTRS

!AA\_SEQUENCE 1.0  
 P1;B44336 - neurotoxin Tx3-2 - spider (Phoneutria nigriventer)

C;Species: Phoneutria nigriventer  
 C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Aug-1998  
 C;Accession: B44336  
 R;Cordeiro M do, N.; de Figueirido, S.G.; Valentim A do, C.; Diniz, C.R.; von Eickstedt, V.R.; Gilroy, J.; Richardson, M.  
 Toxicon 31, 35-42, 1993  
 A;Title: Purification and amino acid sequences of six Tx3 type neurotoxins from the venom of the Brazilian 'armed' spider Phoneutria nigriventer (Keys.).  
 A;Reference number: A44336; MUID:93190315; PMID:8446961  
 A;Accession: B44336  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-34 <COR>  
 A;Experimental source: venom  
 A;Note: sequence extracted from NCBI backbone (NCBIP:126773)  
 C;Superfamily: curatoxin  
 C;Keywords: neurotoxin

B44336 Length: 34 December 22, 2002 19:21 Type: P Check: 2946 ..

1 ACAGLYKKCG KGASPCCEDR PCKCDLAMGN CICK

!AA\_SEQUENCE 1.0  
 F1;B28563 - hemoglobin chain IV - earthworm (*Lumbricus terrestris*) (fragment)  
 C;Species: Lumbricus terrestris (common earthworm)  
 C;Accession: B28563  
 R;Gotoh, T.; Shishikura, F.; Snow, J.W.; Ereifej, K.I.; Vinogradov, S.N.; Walz, D.A.  
 Biochem. J. 241, 441-445, 1987  
 A;Title: Two globin strains in the giant annelid extracellular haemoglobins.  
 A;Reference number: A90337; MUID:87241210; PMID:3595201  
 A;Accession: B28563  
 A;Molecule type: protein  
 A;Residues: 1-22 <GO>  
 C;Superfamily: globin; globin homology  
 C;Keywords: oxygen carrier

B28563 Length: 22 December 22, 2002 19:21 Type: P Check: 9362 ..

1 ADEEESCCSYE DRREVVRHWD DV

!AA\_SEQUENCE 1.0  
 F1;S01808 - hemoglobin AII - tube worm (*Lamellibrachia* sp.) (fragment)  
 C;Species: Lamellibrachia sp.  
 C;Accession: S01808  
 R;Suzuki, T.; Takagi, T.; Ohta, S.  
 Biochem. J. 255, 541-545, 1988  
 A;Title: N-terminal amino acid sequence of the deep sea tube worm haemoglobin remarkably resembles that of annelid haemoglobin.  
 A;Reference number: S01807; MUID:89076216; PMID:320832  
 A;Accession: S01808  
 A;Molecule type: protein  
 A;Residues: 1-24 <SUZ>  
 C;Superfamily: globin; globin homology  
 C;Keywords: oxygen carrier

S01808 Length: 24 December 22, 2002 19:21 Type: P Check: 3406 ..

1 SSNSCTEDR REMOLMWANV WSAQ

!AA\_SEQUENCE 1.0  
 F1;D60894 - gamma crystallin V - bullfrog (fragment)  
 C;Species: Rana catesbeiana (bullfrog)  
 C;Accession: D60894  
 C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Nov-1997  
 R;Chiou, S.H.  
 J. Protein Chem. 7, 527-534, 1988  
 A;Title: The protein sequence homology of gamma-crystallins among major vertebrate classes and their DNA sequence homology to heat-shock protein genes.  
 A;Reference number: A60894; MUID:89351593; PMID:3255376  
 A;Accession: D60894

A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
C;Superfamily: beta-crystallin  
C;Keywords: duplication; eye lens

D60894 Length: 20 December 22, 2002 19:21 Type: P Check: 6311 ..

1 GKVVFYEDRN FGGRSYECSS

!!AA\_SEQUENCE 1.0  
F1;S04621 - gamma-crystallin II - milk shark (fragment)  
C;Species: Rhizoprionodon acutus (milk shark)  
C;Accession: S04621  
R;Chiou, S.H.

FERS Lett. 250, 25-29, 1989  
A;Title: Characterization of gamma-crystallins from eye lenses of shark: closer structural similarity to mammalian than other piscine gamma-crystallins?  
A;Reference number: S04621; MUID:89290026; PMID:2737298  
A;Accession: S04621  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
C;Superfamily: beta-crystallin  
C;Keywords: duplication

S04621 Length: 20 December 22, 2002 19:21 Type: P Check: 5853 ..

1 GKITFYEDRG FGQHGYECSS

!!AA\_SEQUENCE 1.0  
F1;A61570 - gamma-crystallin (total) - siamang (fragment)  
C;Species: Hylobates syndactylus (Siamang)  
C;Date: 25-Oct-1994 #sequence\_revision 06-Jan-1995 #text\_change 21-Nov-1997  
C;Accession: A61570  
R;Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.  
Curr. Eye Res. 7, 107-102, 1988  
A;Title: Biochemical characterization of lens crystallins from three mammalian species.  
A;Reference number: A61570; MUID:89152306; PMID:3229121  
A;Accession: A61570  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
A;Note: this preparation appears to represent the product of several gamma crystallin genes from one specimen  
C;Superfamily: beta-crystallin  
C;Keywords: duplication; eye lens

A61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6107 ..

1 GKITFYEDRG FGQRHYECTT

!!AA\_SEQUENCE 1.0  
F1;B61570 - gamma-crystallin (total) - Formosa sika deer (fragment)  
C;Species: Cervus nippon taivanus (Formosa sika deer)  
C;Accession: Oct-1994 #sequence\_revision 06-Jan-1995 #text\_change 21-Nov-1997  
C;Accession: B61570  
R;Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.  
Curr. Eye Res. 7, 107-102, 1988  
A;Title: Biochemical characterization of lens crystallins from three mammalian species.  
A;Reference number: A61570; MUID:89152306; PMID:3229121  
A;Accession: B61570  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
A;Note: 6-Phe, 14-His, and 15-Cys were also found  
A;Note: this preparation appears to represent the product of several gamma crystallin genes from one specimen  
C;Superfamily: beta-crystallin  
C;Keywords: duplication; eye lens

B61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6068 ..

---

1 GKITFYEDRG FGGRHYECSS

!!AA\_SEQUENCE 1.0  
F1;A60502 - myonexin - northern leopard frog (fragment)  
C;Species: Rana pipiens (northern leopard frog)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C;Accession: A60502; A33087  
R;Tidball, J.G.  
Dev. Biol. 142, 103-114, 1990  
A;Title: Myonexin: an 80-kDa glycoprotein that binds fibronectin and is located at embryonic myotendinous junctions.  
A;Reference number: A60502; MUID:91032554; PMID:2146177  
A;Accession: A60502  
A;Molecule type: protein  
A;Residues: 1-25 <TID>  
C;Comment: This protein resembles calsequestrin in sequence but differs in distribution and physical properties.  
C;Superfamily: calsequestrin  
C;Keywords: fibronectin binding; glycoprotein; skeletal muscle

A60502 Length: 25 December 22, 2002 19:21 Type: P Check: 4514 ..

1 EDGFDFPEYD GDRVFLIST KNYKA

!!AA\_SEQUENCE 1.0  
F1;E33208 - calreticulin, uterine - rabbit (fragment)  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 31-Jul-1991 #sequence\_revision 31-Jul-1991 #text\_change 01-Mar-1996  
C;Accession: E33208; F33208  
R;Milner, R.E.; Baksh, S.; Shevchenko, C.; Carpenter, M.R.; Smillie, L.; Vance, J.E.; Opas, M.; Michalak, M.  
J. Biol. Chem. 266, 7155-7165, 1991  
A;Title: Calreticulin, and not calsequestrin, is the major calcium binding protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.  
A;Reference number: A33208; MUID:91201375; PMID:2016321  
A;Accession: E33208  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-29 <MLN>  
A;Note: uterine form  
A;Accession: F33208  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <ML2>  
A;Note: hepatic form  
C;Superfamily: calreticulin

E33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2821 ..

1 EPVVYFKEQF LDGDSWMDRW IESKHKSDF

!!AA\_SEQUENCE 1.0  
F1;C33208 - calreticulin, slow twitch skeletal muscle - rabbit (fragment)  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 31-Jul-1991 #sequence\_revision 31-Jul-1991 #text\_change 07-Feb-1997  
C;Accession: C33208  
R;Milner, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance, J.E.; Opas, M.; Michalak, M.  
J. Biol. Chem. 266, 7155-7165, 1991  
A;Title: Calreticulin, and not calsequestrin, is the major calcium binding protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.  
A;Reference number: A33208; MUID:91201375; PMID:2016321  
A;Accession: C33208  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-29 <ML>  
C;Superfamily: calreticulin  
C;Keywords: skeletal muscle

C33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2871 ..

!AA\_SEQUENCE 1.0  
F1;A33434 - calcium-binding protein - dog (fragment)  
C;Species: Canis lupus familiaris (dog)  
C;Date: 21-Feb-1990 #sequence\_revision 21-Feb-1990 #text\_change 16-Feb-1997  
C;Accession: A33434  
R;Collins, J.H.; Xie, Z.; Alderson-Hang, B.H.; Treves, S.; Volpe, P.  
Blochem. Biophys. Res. Commun. 164, 575-579, 1989  
A;Title: Sequence homology of a canine brain calcium-binding protein with calreolin and the human Ro/SS-A antigen.  
A;Reference number: A33434; MUID:90026437; PMID:2803321  
A;Accession: A33434  
A;Molecule type: protein  
A;Residues: 1-24 <C0L2>  
C;Keywords: calcium binding  
A;Status: preliminary  
C;Superfamily: calreticulin  
C;Keywords: calcium binding

A33434 Length: 24 December 22, 2002 19:21 Type: P Check: 2817 ..  
1 EPAIYFKEQF LDGGDGXTDRX IESK

!AA\_SEQUENCE 1.0  
F1;B41481 - virulence-associated protein (vira 5' region) - Salmonella typhimurium (fragment)  
C;Species: Salmonella typhimurium  
C;Date: 10-Apr-1992 #sequence\_revision 10-Apr-1992 #text\_change 20-Aug-1999  
C;Accession: B41481  
R;Gulig, P.A.; Chiodo, V.A.  
Infect. Immun. 58, 2651-2658, 1990  
A;Title: Genetic analysis of the Salmonella typhimurium virulence plasmid gene encoding the 28,000-molecular-weight protein.  
A;Reference number: A41481; MUID:90316693; PMID:2164511  
A;Accession: B41481  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-36 <GUL>  
A;Cross-references: GB:Me4295; GB:M33927; NID:9154440; PIDN:AAA27272.1; PMID:9154441  
C;Species family: virulence-associated Protein svyB  
C;Keywords: DNA binding; transcription regulation

B41481 Length: 36 December 22, 2002 19:21 Type: P Check: 1642 ..  
1 GSODDFASOLs KRLSLDDRTA DTNRKRIIN MRVLNS

!AA\_SEQUENCE 1.0  
F1;S07394 - DNA-Invertase - phage D108 (fragment)  
C;Species: Phage D108  
A;Note: host Escherichia coli  
C;Accession: S07394  
R;Szatmari, G.B.; Lapointe, M.; Dubow, M.S.  
Nucleic Acids Res. 15, 6691-6704, 1987  
A;Title: The right end of transposon bacteriophage D108 contains a 520 base pair protein-encoding sequence not present in bacteriophage Mu.  
A;Reference number: S07394; MUID:87316928; PMID:2957646  
A;Accession: S07394  
A;Molecule type: DNA  
A;Residues: 1-22 <SZA>  
A;Cross-references: EMBL:X05926; NID:914918; PIDN:CAA29365.1; PMID:91364174  
C;Genetics:  
A;Gene: gin  
C;Superfamily: transposase repressor  
C;Keywords: DNA binding; DNA integration; DNA recombination  
S07394 Length: 22 December 22, 2002 19:21 Type: P Check: 9004 ..  
1 YKKHPAKRTH IENDDRINQI DR

!AA\_SEQUENCE 1.0  
F1;S35552 - vesicle-associated membrane protein 1 - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999  
C;Accession: S35554  
R;Patarello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.  
Nature 364, 581-582, 1993  
A;Title: Neurotransmission and secretion.  
A;Reference number: S35552; MUID:93354436; PMID:8350916  
A;Accession: S35553  
A;Molecule type: DNA  
A;Residues: 1-32 <PAT>  
C;Superfamily: synaptobrevin

S35552 Length: 32 December 22, 2002 19:21 Type: P Check: 9008 ..  
1 ERDOKLSELD DRADALQAGA SFVFEAAKL KR

!AA\_SEQUENCE 1.0  
F1;S35553 - vesicle-associated membrane protein 1 - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 20-May-1994 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C;Accession: S35553  
R;Patarello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.  
Nature 364, 581-582, 1993  
A;Title: Neurotransmission and secretion.  
A;Reference number: S35552; MUID:93354436; PMID:8350916  
A;Accession: S35553  
A;Molecule type: DNA  
A;Residues: 1-32 <PAT>  
C;Superfamily: synaptobrevin

S35553 Length: 32 December 22, 2002 19:21 Type: P Check: 9118 ..  
1 ERDOKLSELD DRADALQAGA SFVFEAAKL KR

!AA\_SEQUENCE 1.0  
F1;S35555 - vesicle-associated membrane protein 2 - chicken (fragment)  
C;Species: Gallus gallus (chicken)  
C;Date: 10-Dec-1993 #sequence\_revision 26-Jul-1996 #text\_change 17-Mar-1999  
C;Accession: S35555  
R;Patarello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.  
Nature 364, 581-582, 1993  
A;Title: Neurotransmission and secretion.  
A;Reference number: S35552; MUID:93354436; PMID:8350916  
A;Accession: S35555  
A;Molecule type: DNA  
A;Residues: 1-32 <PAT>  
C;Superfamily: synaptobrevin  
C;Keywords: membrane trafficking; transmembrane protein

S35555 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..  
1 ERDOKLSELD DRADALQAGA SFVFEAAKL KR

!AA\_SEQUENCE 1.0  
F1;S35554 - vesicle-associated membrane protein 2 - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999  
C;Accession: S35554  
R;Patarello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.  
Nature 364, 581-582, 1993  
A;Title: Neurotransmission and secretion.  
A;Reference number: S35552; MUID:93354436; PMID:8350916  
A;Accession: S35554  
A;Molecule type: DNA  
A;Residues: 1-32 <PAT>  
C;Superfamily: synaptobrevin

S35554 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..  
1 ERDOKLSELD DRADALQAGA SFVFEAAKL KR

!AA\_SEQUENCE 1.0  
F1;E44621 - homeotic protein Hox 10 (clone 10w) - sea lamprey (fragment)

C;Species: Petromyzon marinus (see lamprey)  
C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 15-Oct-1999  
C;Accession: E44621  
R;Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993  
A;Title: Expansion of the Hox gene family and the evolution of chordates.  
A;Reference number: A44616; MUID:9317669; PMID:8101001  
A;Accession: E44621  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: GB:L14895; NID:9290813; PIDN:AAA02530.1; PID:9290814  
C;Superfamily: unassigned homeobox proteins; homeobox homology  
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

E44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9552 ..

---

1 LFPMWYLTRER RLEISRGYLN TDROVKI

!AA\_SEQUENCE 1.0  
FL:F44621 - homeotic protein Hox 10 (clone 10x) - sea lamprey (fragment)

C;Species: Petromyzon marinus (sea lamprey)  
C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 15-Oct-1999  
C;Accession: F44621  
R;Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993  
A;Title: Expansion of the Hox gene family and the evolution of chordates.  
A;Reference number: A44616; MUID:9317669; PMID:8101001  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-27 <PEN>  
A;Accession: GB:L14900; NID:9290819; PIDN:AAA02535.1; PID:9290820  
C;Superfamily: unassigned homeobox proteins; homeobox homology  
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9407 ..

---

1 LFSSMLTRER RLEISHLSL TDROVKI

!AA\_SEQUENCE 1.0  
PI:C84355 - hypothetical protein Vng2049c [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Accession: C84355  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.;  
Lasky, S.R.; Baliga, N.; Thoraisin, V.; Sbrroga, J.; Smartzell, S.; Weir, D.;  
Hall, J.; Dahl, T.A.; Welti, R.; Goo, Y.A.; Leithauer, B.; Keller, K.; Cruz, R.;  
Danson, M.J.; Rough, D.W.; Maddocks, D.G.; Jablonski, P.E.; Krebs, M.P.;  
Argenine, C.M.; Dale, H.; Isenbarger, T.A.; Peck, R.F.; Pohlschrod, M.;  
Spudich, J.L.; Jung, K.H.; Aleman, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.R.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Liang, P.; Riley, M.; Hood, L.; Bassarana, S.  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: C84355  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-32 <STO>  
A;Cross-references: GB:AE004437; NID:910581474; PIDN:AAG20207.1; GSPDB:GN00138  
A;Genes: VNG2049C

C84355 Length: 32 December 22, 2002 19:21 Type: P Check: 9872 ..

---

1 MVRVPVTDAG RAEMRERADR LETTAFAFWRL VD

!AA\_SEQUENCE 1.0  
PI:E6634 - hypothetical protein HP0917 - Helicobacter pylori (strain 26695)  
C;Species: Helicobacter pylori  
C;Accession: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C;Accession: E64634

!AA\_SEQUENCE 1.0  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klein, H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; Mcineney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gooleyne, J.M.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.M.; Venter, J.C.  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: E64634  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-23 <ROM>  
A;Cross-references: GB:AE000601; GB:AE000511; NID:92314051; PIDN:AD07958.1; PID:92314059; TIGR:HP0917

E64634 Length: 23 December 22, 2002 19:21 Type: P Check: 575 ..

---

1 MSPLTPLRNP LTQEDDRFFQEE IIA

!AA\_SEQUENCE 1.0  
PI:PH0858 - MauD protein - Paracoccus denitrificans (fragment)  
C;Species: Paracoccus denitrificans  
C;Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
C;Accession: PH0858  
R;Christoerdov, A.Y.; Boyd, J.; Mathews, F.S.; Lidstrom, M.E.  
Biochem. Biophys. Res. Commun. 184, 1181-1189, 1992  
A;Title: The genetic organization of the mau gene cluster of the facultative autotroph Paracoccus denitrificans.  
A;Reference number: PH0856; MUID:92272706; PMID:1590782  
A;Accession: PH0858  
A;Molecule type: DNA  
A;Residues: 1-23 <HII>  
A;Cross-references: GB:MG0098; NID:9150580; PIDN:AAA25577.1; PID:9150581  
A;Genetics:  
A;Gene: mauD

PH0858 Length: 23 December 22, 2002 19:21 Type: P Check: 807 ..

---

1 LEADRTGFAS LQOYMASRKQ QAA

!AA\_SEQUENCE 1.0  
PI:A61412 - methane monooxygenase (EC 1.14.13.25) protein Y - Methylococcus capsulatus (fragment)  
C;Species: Methylococcus capsulatus  
C;Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Sep-1994  
C;Accession: A61412  
R;Rosenzweig, A.C.; Lipard, S.J.  
Acc. Chem. Res. 27, 220-236, 1994  
A;Title: Determining the structure of a hydroxylase enzyme that catalyzes the conversion of methane to methanol in methanotrophic bacteria.  
A;Reference number: A61412  
A;Accession: A61412  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-27 <ROS>  
C;Keywords: oxidoreductase

A61412 Length: 27 December 22, 2002 19:21 Type: P Check: 8072 ..

---

1 WIEDYASRID FKADRDQIVK AVLAGLK

!AA\_SEQUENCE 1.0  
PI:A82288 - hypothetical protein VC0735 [imported] - Vibrio cholerae (strain NL6961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Accession: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C;Accession: A82288

R; Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwynn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, M.; S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Erauvaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishman, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000  
A; Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A; Reference number: A82035; MUID:20406833; PMID:10952301  
A; Accession: A82288  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-33 <HET>  
A; Cross-references: GB:AE004159; GB:AE003852; NID:99655167; PIDN:AAF93900.1; GSPDS:GN00126; TIGR:VCA735  
A; Experimental source: serogroup O1; strain N16961; biotype El Tor  
C; Generics:  
A; Gene: VCA735  
A; Map position: 1

A82288 Length: 33 December 22, 2002 19:21 Type: P Check: 3787 ..

1 MPVGSSIPCS ISTKLADRKV VSRWMCFFL LW

!!AA\_SEQUENCE 1.0  
P1;B6421 - hypothetical protein VCA0761 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C;Accession: B6421  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwynn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, M.; S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Erauvaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishman, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000  
A; Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A; Reference number: A82035; MUID:20406833; PMID:10952301  
A; Accession: B6421  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-32 <HET>  
A; Cross-references: GB:AE004404; GB:AE003853; NID:99658174; PIDN:AAF96659.1; GSPDB:GN00127; TIGR:VCA761  
A; Experimental source: serogroup O1; strain N16961; biotype El Tor  
C; Generics:  
A; Gene: VCA0761  
A; Map position: 2

B82421 Length: 32 December 22, 2002 19:21 Type: P Check: 1300 ..

1 MRDRILRPSA ICSDRRKQT NFESSDLKQQ WR

!!AA\_SEQUENCE 1.0  
P1;B39089 - hydrogenase (EC 1.18.99.1) 34K chain - Thiocapsa roseopersicina (fragment)  
C;Species: Thiocapsa roseopersicina  
C;Date: 27-Nov-1991 #sequence\_revision 27-Nov-1991 #text\_change 23-Jun-1993  
C;Accession: B39089  
R; Kovacs, K.L.; Tigray, G.; Thanha, L.T.; Lakatos, S.; Kiss, Z.; Bagyinka, C.  
A; Title: Structural rearrangements in active and inactive forms of hydrogenase from Thiocapsa roseopersicina.  
A; Reference number: A39089; MUID:91093297; PMID:1845598  
A; Accession: B39089  
R; Status: preliminary  
A; Molecule type: protein  
A; Residues: 1-20 <KOV>  
C;Keywords: oxidoreductase

B39089 Length: 20 December 22, 2002 19:21 Type: P Check: 6231 ..

1 XXEADRPSV IGLPFQEXTG

!!AA\_SEQUENCE 1.0  
P1;H70954 - hypothetical protein Rv3599c - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Accession: H70954  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekla, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jacobs, K.; Krogh, A.; McLean, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. Nature 393, 537-544, 1998  
A; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.  
A; Reference number: A70500; MUID:98295987; PMID:9634230  
A; Accession: H70954  
A; Status: preliminary; nucleic acid sequence not shown; translation not shown  
A; Molecule type: DNA  
A; Residues: 1-7 <COL>  
A; Cross-references: GB:295557; G3:AL123456; NID:93242276; PIDN:CAR08945.1; PID:ie31647; PID:2113977  
A; Experimental source: strain H37RV  
C; Generics:  
A; Gene: RV3599c

H70954 Length: 27 December 22, 2002 19:21 Type: P Check: 8577 ..

1 KPASSLGTS PAADRDAATH ERREVI

!!AA\_SEQUENCE 1.0  
P1;S14161 - probable drbp-4-keto-6-deoxy-hexose 3,4-isomerase (EC 5.3.99.-) - Saccharopolyspora erythraea (fragment)  
N;Alternate names: 5'K Protein  
C;Species: Saccharopolyspora erythraea  
C;Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 21-Nov-1998  
C;Accession: S14161  
R;Caffrey, P.; Green, B.; Packman, L.C.; Rawlings, B.J.; Staunton, J.; Leadlay, P.F.  
Bur. J. Biochem. 195, 823-830, 1991  
A;Title: An acyl-carrier-protein - thioesterase domain from the 6-deoxyerythronolide B synthase of *Saccharopolyspora erythraea*. High-level production, purification and characterisation in *Escherichia coli*.  
A; Reference number: S14091; MUID:91153324; PMID:1999200  
A; Accession: S14161  
A; Molecule type: protein  
A; Residues: 1-20 <CAF>  
C;Generics:  
A; Gene: erC11  
A; Function:  
A; Description: involved in desosamine biosynthesis  
C;Keywords: antibiotic biosynthesis; intramolecular oxidoreductase; isomerase

S14161 Length: 20 December 22, 2002 19:21 Type: P Check: 5957 ..

1 MTTDRAGLG ROLOMIRGLH

!!AA\_SEQUENCE 1.0  
P1;S66213 - glucose 1-dehydrogenase (EC 1.1.1.47) - Haloferax mediterranei (fragment)  
C;Species: Haloferax mediterranei  
C;Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 30-Oct-1998  
C;Accession: S66213  
R; Botte, M.J.; Pier, C.; Llorca, F.I.; Camacho, M.I.  
FBS Lett. 383, 221-229, 1995  
A;Title: Glucose dehydrogenase from the halophilic archaeon *Haloferax mediterranei*: enzyme purification, characterization and N-terminal sequence.

A; Reference number: S66213; MUID:96198607; PMID:8925901  
A; Accession: S66213  
A; Status: preliminary  
A; Molecule type: protein  
A; Residues: 1-17 <BON>  
C; Keywords: oxidoreductase

S66213 Length: 17 December 22, 2002 19:21 Type: P Check: 1672 ..

1 MKAIAVKRG EDRPVVIE

!AA\_SEQUENCE 1.0  
PI;S23180 - carboxypeptidase - *Sulfolobus solfataricus*  
C;Species: *Sulfolobus solfataricus* 19-Mar-1997 #text\_change 24-Jul-1997  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
R;Colombo, S.; D'Auria, S.; Fusi, P.; Zecca, L.; Raia, C.A.; Tortora, P.  
Eur. J. Biochem. 206, 349-357, 1992  
A;Title: Purification and characterization of a thermostable carboxypeptidase from the extreme thermophilic archaebacterium *Sulfolobus solfataricus*.  
A;Reference number: S23180; MUID:92283259; PMID:1597179  
A;Accession: S23180  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-31 <COL>

S23180 Length: 31 December 22, 2002 19:21 Type: P Check: 7956 ..

1 MDIVEKLND VKEHEDRIIQ IRKTKXENPP L ..

!AA\_SEQUENCE 1.0  
PI;S23936 - hypothetical protein 22, psba 5'-region - Japanese black pine chloroplast  
C;Species: chloroplast *Pinus thunbergiana* (Japanese black pine)  
C;Accession: S23936; T07427  
R;Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Hiratsuka, J.; Shibata, M.; Wakasugi, T.; Sugura, M.  
Mol. Gen. Genet. 242, 206-214, 1992  
A;Title: Chloroplast DNA of black pine retains a residual inverted repeat lacking rRNA genes: nucleotide sequences of trnQ, trnK, psba, trnI and trnH and the absence of rps16  
A;Residues: 1-22 <TSU>  
A;Cross-references: EMBL:D11467; NID:9344007; PID:BA02023.1; PID:9344011; R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugura, M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994  
A;Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*:  
A;Reference number: Z16030; MUID:95024047; PMID:7937893  
A;Accession: T07427  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-22 <WAK>  
A;Cross-references: EMBL:D17510; NID:9529643; PIDN:BAA04307.1; PID:91262588  
C;Genetics:  
A;Genome: chloroplast  
C;Keywords: chloroplast

S23936 Length: 22 December 22, 2002 19:21 Type: P Check: 9119 ..

1 MEYLTDRSI ECGIYLLKIE SI ..

!AA\_SEQUENCE 1.0  
PI;T10123 - probable catalase (EC 1.11.1.6) - cucumber (fragment)  
C;Species: *Cucumis sativus* (cucumber)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C;Accession: T10123  
R;Toyama, T.; Teramoto, H.; Takeba, G.; Tsuji, H.  
Plant Cell Physiol. 36, 1349-1359, 1995

A;Title: Cytokinin induces a rapid decrease in the levels of mRNAs for catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and other unidentified proteins in etiolated cotyledons of cucumber.  
A;Reference number: Z16946; MUID:96104306; PMID:8564304  
A;Accession: T10123  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-23 <TOY>  
A;Cross-references: EMBL:D63385; NID:91199475; PIDN:BAA09701.1; PID:91199476  
A;Experimental source: seedling; cotyledons  
C;Keywords: oxidoreductase

T10123 Length: 23 December 22, 2002 19:21 Type: P Check: 1444 ..

1 NGSQADRSVG QKLAPHLNVR PSI ..

!AA\_SEQUENCE 1.0  
PI;P0212 - 29K protein 4228 - rice (strain Nihonbare) (fragment)  
C;Species: *Oryza sativa* (rice)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995  
C;Accession: P0212  
R;Tsudita, A.; Miyatake, N.  
Submitted to JPIID, April 1993  
A;Reference number: FS2028  
A;Accession: P0212  
A;Molecule type: protein  
A;Residues: 1-15 <TSU>  
A;Experimental source: germ  
C;Comment: molecular weight 29K, pi 6.1.

PS0212 Length: 15 December 22, 2002 19:21 Type: P Check: 8963 ..

1 SPADDRRDVG DRYAD ..

!AA\_SEQUENCE 1.0  
PI;B60698 - trichocyst protein 27 - *Paramecium tetraurelia* (fragment)  
C;Species: *Paramecium tetraurelia*  
C;Accession: B60698  
R;Tindall, S.H.; Devito, L.D.; Nelson, D.L.  
J. Cell Sci. 92, 441-447, 1989  
A;Title: Biochemical characterization of the proteins of paramecium secretory granules.  
A;Residues: 1-29 <TSU>  
C;Comment: This protein was purified from the extruded matrix (contents) of trichocysts (secretory granules).  
A;Genetic code: SGC5  
A;Molecule type: protein  
C;Keywords: extracellular protein

B60698 Length: 29 December 22, 2002 19:21 Type: P Check: 2526 ..

1 DPDLRDLSTL TDLEDRYVAE QKEDDAKNG ..

!AA\_SEQUENCE 1.0  
PI;S19614 - globin - polychaete (*Eudistylia vancouveri*) (fragment)  
N;Alternate names: chlorocruorin  
C;Species: *Eudistylia vancouveri*  
C;Accession: S19614  
C;Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 05-Dec-1998  
R;Qabbar, O.H.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Tinkovich, R.; Wall, J.S.; J. Mol. Biol. 222, 1109-1129, 1991  
A;Title: Hierarchy of globin complexes. The quaternary structure of the extracellular chlorocruorin of *Eudistylia vancouveri*.  
A;Reference number: S19532; MUID:9210633; PMID:1762147  
A;Accession: S19614  
A;Molecule type: protein  
A;Residues: 1-17 <QAB>  
A;Experimental source: plume

C;Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dodecamers are linked into bilayer structure by Ca(2+) and heme-deficient chimeric globin chains  
C;Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

S19614 Length: 17 December 22, 2002 19:21 Type: P Check: 1820 ..

1 GMKXXSMDR KTVLADW

! ! ! AA\_SEQUENCE 1.0  
F1;A60741 - insulin-like growth factor-binding protein He39L - human (fragment)  
C;Species: Homo sapiens (man)  
C;Accession: A60741  
R;Forbes, B.; Ballard, F.J.; Wallace, J.C.  
J; Endocrinol. 126, 497-506, 1990  
A;Title: An insulin-like growth factor-binding protein purified from medium conditioned by a human lung fibroblast cell line (He(39)L) has a novel N-terminal sequence.  
A;Reference number: A60741; MUID:91011238; PMID:1698907  
A;Accession: A60741  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-25 <FOR>

A60741 Length: 25 December 22, 2002 19:21 Type: P Check: 4352 ..

1 IAPGPGSGVQ AGXPXGXVEE EDRGG

! ! ! AA\_SEQUENCE 1.0  
F1;B28457 - proteoglycan II, bone - human (tentative sequence) (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 19-May-1989 #sequence\_revision 19-May-1989 #text\_change 31-Mar-2000  
C;Accession: B28457  
R;Fisher, L.W.; Hawkins, G.R.; Tuross, N.; Termine, J.D.  
J; Biol. Chem. 262, 9702-9708, 1987  
A;Title: Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone.  
A;Reference number: A92656; MUID:87250639; PMID:3597437  
A;Accession: B28457  
A;Molecule type: protein  
A;Residues: 1-19 <FIS>

B28457 Length: 19 December 22, 2002 19:21 Type: P Check: 4360 ..

1 DEAXGIAAPEV PDDDRFFEPS

! ! ! AA\_SEQUENCE 1.0  
F1;C5037 - splicing regulatory protein SWAP homolog (alternatively spliced, clone pFL2) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 28-Feb-1997  
C;Accession: C5037  
R;Dennie, F.; Narayatis, R.  
J; Biol. Chem. 269, 16170-16179, 1994  
A;Title: Conservation of regulated alternative splicing and identification of functional domains in vertebrate homologs to the Drosophila splicing regulator, suppressor-of-white-apricot.  
A;Reference number: A54037; MUID:91266805; PMID:8206918  
A;Accession: C54037  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-29 <DEN>  
A;Source: liver  
A;Note: sequential source extracted from NCBI backbone (NCIP:149928)  
C54037 Length: 29 December 22, 2002 19:21 Type: P Check: 1565 ..

1 ! ! ! AA\_SEQUENCE 1.0  
F1;A48845 - sterol regulatory element 1 binding protein (alternatively spliced,

clone pcY22) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C;Accession: A48845  
R;Yokoyama, C.; Wang, X.; Briggs, M.R.; Admon, A.; Wu, J.; Hua, X.; Goldstein, J.L.; Brown, M.S.  
Cell 75, 187-197, 1993  
A;Title: SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene.  
A;Reference number: A48845; MUID:9406541; PMID:8402897  
A;Accession: A48845  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-23 <OK>  
A;Cross-references: GB:S66168;  
A;Experimental source: HeLa S3 cells  
A;Note: Sequence extracted from NCBI backbone (NCBIN:138566, NCBIN:138568, NCBIPI:18870)

A48845 Length: 23 December 22, 2002 19:21 Type: P Check: 9975 ..

1 RGRGANGTDA P RAGADRGAMD CTF

! ! ! AA\_SEQUENCE 1.0  
F1;S22228 - vitronectin - dog (fragment)  
C;Species: Canis lupus familiaris (dog)  
C;Date: 22-Nov-1993 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997  
C;Accession: S22228  
R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori, H.; Matsumoto, I.; Seno, N.; Hayashi, M.  
Biochim. Biophys. Acta 1120, 1-10, 1992  
A;Title: Vitronectin diversity in evolution but uniformity in ligand binding and size of the core polypeptide.  
A;Reference number: S21768; MUID:92207982; PMID:1372829  
A;Accession: S22228  
A;Molecule type: protein  
A;Residues: 1-25 <MAK>

S22228 Length: 25 December 22, 2002 19:21 Type: P Check: 4810 ..

1 AQESXKGRTV EGFNADRKKQ QDELX

! ! ! AA\_SEQUENCE 1.0  
F1;A26393 - annexin 36K chain - pig (fragment)  
C;Species: Sus scrofa domesticus (domestic pig)  
C;Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 18-Jun-1993  
C;Accession: A26393  
R;Geisow, M.J.; Fritschke, U.; Hexham, J.M.; Dash, B.; Johnson, T.  
Nature 330, 636-638, 1986  
A;Title: A consensus amino-acid sequence repeat in Torpedo and mammalian Ca(2+)-dependent membrane-binding proteins.  
A;Reference number: A93379; MUID:86203621; PMID:2422556  
A;Accession: A26393  
A;Molecule type: protein  
A;Residues: 1-16 <GET>

A26393 Length: 16 December 22, 2002 19:21 Type: P Check: 7 ..

1 MLGLCTDEDR LILIL

! ! ! AA\_SEQUENCE 1.0  
F1;S68260 - hypothetical protein gadd7.1 - long-tailed hamster  
C;Species: Cricetus longicaudatus (long-tailed hamster)  
C;Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C;Accession: S68260  
R;Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.  
Nucleic Acids Res. 24, 1589-1593, 1996  
A;Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth, but lacks a protein product.  
A;Reference number: S68260; MUID:96211359; PMID:8649973  
A;Accession: S68260  
A;Status: preliminary  
A;Molecule type: mRNA

A;Residues: 1-38 <HOL>  
A;Cross-references: EMBL:140430  
S68260 Length: 38 December 22, 2002 19:21 Type: P Check: 6375 ..

1 MPSEFSSK QLQRAQMED RGPHPSEFI TGNLKLTK  
! AA\_SEQUENCE 1.0  
PI:A35678 - hypothetical protein (proenkephalin 5' region) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 05-Nov-1999  
C;Accession: A35678  
R;Kilpatrick, D.L.; Zinn, S.A.; Fitzgerald, M.; Higuchi, H.; Sabol, S.L.;  
Mol. Cell. Biol. 10, 3717-3726, 1990  
A;Title: Transcription of the rat and mouse proenkephalin genes is initiated at distinct sites in spermatogenic and somatic cells.  
A;Reference number: A35678; MUID:90287163; PMID:2155920  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-18 <KIL>  
A;Cross-references: GB:M55181; NID:9201032; PIDN:AAA40127.1; PID:9201033  
A35678 Length: 18 December 22, 2002 19:21 Type: P Check: 3193 ..

1 MSGKDKPSW EDRIPPGR  
! AA\_SEQUENCE 1.0  
PI:PH0780 - T-cell receptor alpha chain (Cl1) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Accession: PH0780 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide antigen specific repertoire.  
A;Reference number: PH0780 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
A;Experimental Source: T lymphocyte  
A;Accession: PH0781  
A;Molecule type: mRNA  
A;Residues: 1-15 <CAI>  
A;Cross-references: EMBL:X60915  
A;Experimental source: T lymphocyte  
A;Accession: PH0781  
A;Molecule type: mRNA  
A;Residues: 1-15 <CA2>  
A;Cross-references: EMBL:X60880  
A;Experimental source: T lymphocyte  
C;Keywords: T-cell receptor  
PH0806 Length: 15 December 22, 2002 19:21 Type: P Check: 8864 ..

1 CALSDQGGAD RLTFG  
! AA\_SEQUENCE 1.0  
PI:S03505 - T-cell receptor alpha chain J region (80) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 30-May-1997  
C;Accession: S03505  
R;Winoto, A.; Moisness, S.; Hood, L.  
Nature 316, 832-835, 1985  
A;Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.  
A;Reference number: S03503; MUID:85296332; PMID:2933908  
A;Accession: S03505  
A;Molecule type: DNA  
A;Residues: 1-20 <WIN>  
A;Cross-references: EMBL:X02859  
A;Note: this sequence was determined from the germline gene  
C;Keywords: T-cell receptor  
PH0780 Length: 15 December 22, 2002 19:21 Type: P Check: 8887 ..

1 CASETGGAD RLTFG  
! AA\_SEQUENCE 1.0  
PI:PH0794 - T-cell receptor alpha chain (K1 v-alpha-3 . PHDS58) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PH0794 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.  
A;Reference number: PH0746; MUID:92078846; PMID:1836010  
A;Accession: PH0794  
A;Molecule type: mRNA  
A;Residues: 1-17 <CAS>  
A;Cross-references: EMBL:X60899  
A;Experimental source: T lymphocyte  
C;Keywords: T-cell receptor  
PH0794 Length: 17 December 22, 2002 19:21 Type: P Check: 1531 ..

1 CAVSMNYRG ADDRATFG  
! AA\_SEQUENCE 1.0  
PI:PT0212 - T-cell receptor beta chain V-D-J region (126-1AI) - mouse (fragment)  
C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C;Accession: PT0547  
 R;Freaney, A.J.  
 J.;Exp. Med. 174, 115-124, 1991  
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A;Reference number: PR0509; MUID:91277601; PMID:1711558  
 A;Accession: PT0547  
 A;Status: translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-8 <FEE>  
 A;Experimental source: day 18 fetal thymus, strain BALB/c  
 C;Keywords: T-cell receptor

PT0547 Length: 8 December 22, 2002 19:21 Type: P Check: 2627 ..

1 ASSDADRG

---

! !AA\_SEQUENCE 1.0  
 F1;PT0576 - T-cell receptor beta chain V-D-J region (140-1AL) - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C;Accession: PT0576  
 R;Feeney, A.J. 1991  
 J. Exp. Med. 174, 115-124, 1991  
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A;Reference number: PR0509; MUID:91277601; PMID:1711558  
 A;Status: translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-7 <FEE>  
 A;Experimental source: day 18 fetal thymus, strain BALB/c  
 C;Keywords: T-cell receptor

PT0676 Length: 7 December 22, 2002 19:21 Type: P Check: 2049 ..

1 ASGEDDRG

---

! !AA\_SEQUENCE 1.0  
 F1;PT0576 - T-cell receptor beta chain V-D-J region (141-1G) - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C;Accession: PT0576  
 R;Feeney, A.J. 1991  
 J. Exp. Med. 174, 115-124, 1991  
 A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A;Reference number: PR0509; MUID:91277601; PMID:1711558  
 A;Accession: PT0576  
 A;Status: translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-7 <FEE>  
 A;Experimental source: day 19 fetal thymus, strain BALB/c  
 C;Keywords: T-cell receptor

PT0676 Length: 7 December 22, 2002 19:21 Type: P Check: 2049 ..

---

1 ASSDADRG

---

! !AA\_SEQUENCE 1.0  
 F1;PT0576 - T-cell receptor beta chain V-D-J region (6R2) - mouse (fragment)  
 C;Species: Mus musculus (house mouse)  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 30-May-1997  
 C;Accession: PR0366  
 R;Lehmann, P.V.; Drexler, K.; Tary-Lehmann, M.; Falcioni, F.; Hirttenbach, U.; Nagy, Z.A.  
 J. Exp. Med. 173, 333-341, 1991  
 A;Title: Graft-versus-host resistance induced by class II major histocompatibility complex-specific T cell clones.  
 A;Reference number: PR0360; MUID:91108330; PMID:1824856  
 A;Accession: PR0366  
 A;Molecule type: mRNA

---

A;Residues: 1-28 <LEH>  
 C;Keywords: T-cell receptor

PT0366 Length: 28 December 22, 2002 19:21 Type: P Check: 480 ..

1 LYFCASSEDR NNQURELFRG LDPSVLED

---

! !AA\_SEQUENCE 1.0  
 F1;A6592 - lactase-phlorizin hydrolase, 200K isoform - rat (fragment)  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 01-Nov-1996  
 C;Accession: A6592  
 R;Dudley, M.A.; Hatchey, D.L.; Quaroni, A.; Hutchens, T.W.; Nichols, B.L.; Rosenberger, J.; Perkinson, J.S.; Cook, G.; Reeds, P.J.  
 A;Title: In vivo sucrase-isomaltase and lactase-phlorizin hydrolase turnover in the fed adult rat.  
 A;Reference number: A46592; MUID:93299888; PMID:8514793  
 A;Accession: A46592  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-17 <PDB>  
 A;Note: Sequence extracted from NCBI backbone (NCBIPR:134559)  
 C;Keywords: carbohydrate digestion; intestine

A46592 Length: 17 December 22, 2002 19:21 Type: P Check: 1363 ..

1 DWEDDRNFIAA GPLNL

---

! !AA\_SEQUENCE 1.0  
 F1;S78414 - ribosomal protein RL25, mitochondrial [validated] - rat (tentative sequence) (fragment)  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jul-2000  
 C;Accession: S78414  
 R;Goldschmidt-Reisin, S.; Graack, H.R.; submitted to the Protein Sequence Database, February 1998  
 A;Reference number: S78411  
 A;Accession: S78414  
 A;Molecule type: protein  
 A;Residues: 1-26 <GO>  
 C;Keywords: mitochondrion; protein biosynthesis; ribosome

S78414 Length: 26 December 22, 2002 19:21 Type: P Check: 6085 ..

1 ISRREKEKKN IVPDQLIDGE DRRDAE

---

! !AA\_SEQUENCE 1.0  
 F1;S2227 - Vitronectin - guinea pig (fragment)  
 C;Species: Cavia porcellus (guinea pig)  
 C;Date: 22-Nov-1993 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997  
 C;Accession: S2227  
 R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori, H.; Matsumoto, I.; Seno, N.; Hayashi, M.  
 Biochim. Biophys. Acta 1120, 1-10, 1992  
 A;Title: Vitronectin diversity in evolution but uniformity in ligand binding and size of the core polypeptide.  
 A;Reference number: S21768; MUID:92207982; PMID:1372829  
 A;Accession: S2227  
 A;Molecule type: protein  
 A;Residues: 1-25 <NAK>

S2227 Length: 25 December 22, 2002 19:21 Type: P Check: 5585 ..

1 XIXSXKGRXT EGFNADRKKQ XXELX

---

! !AA\_SEQUENCE 1.0  
 F1;A54226 - light-harvesting protein B-830 alpha-1 chain - Chromatium purpureum (fragment)  
 C;Species: Chromatium purpureum  
 C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-Jan-1996  
 C;Accession: A54226

R; Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.  
 Biochemistry 33, 2178-2184, 1994  
 A; Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.  
 A; Reference number: A54226; MUID:94162224; PMID:8117674  
 A; Accession: A54226  
 A; Status: preliminary  
 A; Molecule type: protein  
 A; Residues: 1-26 <KERS>  
 C; Keywords: antenna complex; light-harvesting polypeptide

A54226 Length: 26 December 22, 2002 19:21 Type: P Check: 6170 ..

1 MKVPMVMADE NAKLNPNEDD RKKFFV

---

! IAA\_SEQUENCE 1.0  
 F1;B54226 - light-harvesting protein B-830 alpha-2 chain - Chromatium purpuratum (fragment)  
 C;Species: Chromatium purpuratum  
 C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996  
 C;Accession: B54226  
 R;Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.  
 Biochemistry 33, 2178-2184, 1994  
 A;Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.  
 A;Reference number: A54226; MUID:94162224; PMID:8117674  
 A;Accession: B54226  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-23 <KERS>  
 C;Keywords: antenna complex; light-harvesting polypeptide

B54226 Length: 23 December 22, 2002 19:21 Type: P Check: 634 ..

1 MQPVVMLADK NAKLNHPPEDD RKR

---

! IAA\_SEQUENCE 1.0  
 F1;T46593 - phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)  
 C;Species: Mycobacterium marinum  
 C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
 C;Accession: T46593  
 R;Ramakrishnan, L.; Tran, H.T.; Fedderspiel, N.A.; Falkow, S.  
 J. Bacteriol. 179, 5862-5868, 1997  
 A;Title: A crtB homolog essential for photochromogenicity in Mycobacterium marinum: isolation, characterization, and gene disruption via homologous recombination.  
 A;Reference number: 223096; MUID:97440138; PMID:9294446  
 A;Accession: T46593  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-38 <RAM>  
 A;Cross-references: EMBL:U92075; NID:91928930; PIDN:AAB71427; PID:91928931  
 A;Experimental source: strain M  
 C;Genetics:  
 A;Gene: crtI

T46593 Length: 38 December 22, 2002 19:21 Type: P Check: 6757 ..

1 VPGVGVPITL ISGRLLAIDRI TGNTVRSIRH LDIIKAQLS

---

! IAA\_SEQUENCE 1.0  
 P1;H85575 - hypothetical protein z0899 [imported] - Escherichia coli (strain O157:H7, sustrain ED93)

C;Species: Escherichia coli  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C;Accession: H85575  
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimambra, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.

---

Nature 409, 529-533, 2001  
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A;Reference number: A85480; MUID:21074935; PMID:11206551  
 A;Accession: H85575  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-13 <SPO>  
 A;Cross-references: GB:AE005174; NID:912513665; PIDN:AAG555068; 1; GSPDB:GN00145; UWGZ:Z0899  
 A;Experimental source: strain O157:H7, substrain ED93  
 C;Genetics:  
 A;Gene: 20899

H85575 Length: 13 December 22, 2002 19:21 Type: P Check: 6940 ..

1 MSTDRKPWML LFH

---

! IAA\_SEQUENCE 1.0  
 F1;PA0041 - plastocuonol-plastocyanin reductase (EC 1.10.99.1) - Arabidopsis thaliana (fragment)  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 03-Jun-2002  
 C;Accession: PA0041  
 R;Kano, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
 submitted to JPIB, July 1994  
 A;Description: Separation and characterization of Arabidopsis proteins by two-dimensional gel electrophoresis.  
 A;Reference number: PA0001  
 A;Accession: PA0041  
 A;Molecule type: protein  
 A;Residues: 1-15 <KAM>  
 A;Experimental source: leaf  
 C;Keywords: oxidoreductase

PA0041 Length: 15 December 22, 2002 19:21 Type: P Check: 9117 ..

1 ASSIPADRVF DMERK

! IAA\_SEQUENCE 1.0 STANDARD; PRT; 27 AA.

ID AL20\_CARMA STANDARD; PRT; 27 AA.

AC P81823; DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isozyme M) (Fragment).

DT 01-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE carcinotatin 20.

OS *Carcinus maenas* (Common shore crab) (Green crab).

OC Malacostraca; Eumalacostraca; Crustacea; Pancrustacea; Decapoda; Pleocyemata; Brachyura; Eurbrachyura; Portunoidae; Portunidae; Carcinus.

OC NCBI\_TAXID=6759;

RN [1]

RP SEQUENCE.

RC TISSUE=Cerebral ganglion, and Thoracic ganglion;

RX MEDLINE=99121193; PubMed=941295;

RA Dueve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P., Thorpe A.;

RT Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab *Carcinus maenas*.<sup>n</sup>

RL Eur. J. Biochem. 250:727-734(1997).

CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.

CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

KW Neuropeptide; Amidation; Multigene family.

FT MOD\_RES 27 27 AMIDATION (POTENTIAL);

SQ SEQUENCE 27 AA; 3152 MW; 597A6901965PE987 CRC64;

A120\_CARMA Length: 27 December 22, 2002 19:30 Type: P Check: 9046 ..

1 GYEDEDEDRPF YFALGLGKRP RYNGCL

! IAA\_SEQUENCE 1.0 STANDARD; PRT; 26 AA.

ID AMD1\_CHICK STANDARD; PRT; 26 AA.

AC P81073; DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isoform M) (Fragment).

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isoform M) (Fragment).

OS Gallus gallus (Chicken).

OC Archosauria: Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OC NCBI\_TAXID=9031;

RN [1]

RP SEQUENCE.

RC TISSUE=Breast;

RX MEDLINE=97469365; PubMed=9114497;

RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.; "AMP-deaminases from chicken and rabbit muscle: partial primary sequences of homologous 17-kDa CNBr fragments: autorecognition by rabbit anti-chicken AMPD"; Comp. Biochem. Physiol. 116B:371-377 (1997).

CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY METABOLISM.

CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).

CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.

CC -!- SUBUNIT: HOMOTETRAMER.

CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE, EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3 IS FOUND IN ERYTHROCYTES.

CC DR InterPro:IPR01365; AMP deaminase.

DR PROSITE: PS00485; A\_DEAMINASE; PARTIAL.

KW Hydrolase; Nucleotide metabolism; Multigene family.

FT NON\_TER 26 25

SQ SEQUENCE 26 AA; 3169 MW; B022467EACBB6E75 CRC64;

AMD1\_RABIT Length: 26 December 22, 2002 19:30 Type: P Check: 7826 ..

1 MNQKHLRLFI KKSYYQVDAKR WYISK

! IAA\_SEQUENCE 1.0 STANDARD; PRT; 22 AA.

ID DMY\_BPD10 STANDARD; PRT; 22 AA.

AC Q38190; DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isoform M) (Fragment).

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE DNA-Invertase (Fragment).

GN GIN.

OS Bacteriophage D108.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae; Mu-like viruses.

OC NCBI\_TAXID=10671;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=8731628; PubMed=2257646;

RA Szatmari G.B., Lapointe M., Dubow M.S.; "The right end of the transposable bacteriophage D108 contains a 520 base pair protein-encoding sequence not present in bacteriophage Mu"; Nucleic Acids Res. 15:6691-6703(1987).

RL InterPro:IPR001365; AMP deaminase.

DR PROSITE:PS00405; A\_DEAMINASE; PARTIAL.

KW Hydrolase; Nucleotide metabolism; Multigene family.

FT NON\_TER 1 1

SQ SEQUENCE 26 AA; 3195 MW; B03E296D63BB6E75 CRC64;

AMDI\_CHICK Length: 26 December 22, 2002 19:30 Type: P Check: 7006 ..

1 MNOKHLRFI KKSYYQVDAKR WYISK

! IAA\_SEQUENCE 1.0 STANDARD; PRT; 26 AA.

ID AMD1\_RABIT PRT; 26 AA.

AC P81072;

This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).



**Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment). FGB.**  
 OS Cervus elaphus (Red deer), and  
 OC Cervus elaphus nelsoni (American elk).  
 OC Euarcto; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
 OC Cervidae; Cervinae; Cervus.  
 NCBI\_TaxID=9860; 9864;  
 RN [1]  
 RP SEQUENCE.  
 SPECIES=C. elaphus;  
 RL Blomback B., Blomback M., Grondahl N.J.;  
 RL "Studies on fibrinopeptides from mammals.",  
 Acta Chem. Scand. 19:1789-1791(1965).  
 RN [2]  
 SPECIES=C. e. nelsoni;  
 RC Micro G.A., Doolittle R.F.;  
 RC "Amino acid sequence studies on artiodactyl fibrinopeptides.",  
 Arch. Biochem. Biophys. 122:674-684(1967).  
 RL -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLIMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 THROMBIN, WHICH CLEAVES FIBRINOPETIDES A AND B FROM ALPHA & BETA  
 CHAINS, AND, THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 InterPro: IPRO02181; Fibrinogen\_C.  
 PROSITE: PS00314; FIBRIN\_AGC\_DOMAIN; PARTIAL.  
 DR PROSTKE; PS00314; FIBRIN\_AGC\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation.  
 FT PEPTIDE 1 21 P  
 RT MOD. RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD. RES 6 6 SULFATION.  
 FT NON\_TER 21 21 P  
 SEQUENCE 21 AA: 2558 MW: FCEET45D98931627 CRC64;  
 FIBB\_CEREL Length: 21 December 22, 2002 19:30 Type: P Check: 6821 ..  
 1 OHSTYDEEE EDRAKLHDDA R  
 !!!A\_ SEQUENCE 1.0  
 ID FIBB\_FELCA STANDARD: PRT; 20 AA.  
 AC P14465;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Felis silvestris catus (Cat).  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE.  
 RL Blomback B., Blomback M., Grondahl N.J.;  
 RL "Studies on fibrinopeptides from mammals.",  
 Acta Chem. Scand. 19:1789-1791(1965).  
 RN [2]  
 SPECIES=L. vicugna;  
 RC Blomback B., Blomback M., Grondahl N.J.;  
 RC "Studies on fibrinopeptides from mammals.",  
 Arch. Biochem. Biophys. 122:674-684(1967).  
 RL DR Dromedary camel; mule deer; and cape buffalo.;"  
 RL Arch. Biochem. Biophys. 118:456-467(1967).  
 RN [3]  
 RP SEQUENCE.  
 RC Micro G.A., Doolittle R.F.;  
 RC "Amino acid sequence studies on artiodactyl fibrinopeptides. I."  
 RT DR Dromedary camel; mule deer; and cape buffalo.;"  
 RT Arch. Biochem. Biophys. 118:456-467(1967).  
 RN DR InterPro: IPRO02181; Fibrinogen\_C.  
 DR PROSITE: PS00514; FIBRIN\_AGC\_DOMAIN; PARTIAL.  
 FT KW Blood coagulation; Plasma; Sulfation.  
 FT PEPTIDE 1 19 P  
 FT MOD. RES 4 4 PYRROLIDONE CARBOXYLIC ACID.  
 FT NON\_TER 19 19 P  
 SQ SEQUENCE 19 AA: 2295 MW: E7E6B6100568638 CRC64;  
 FIBB\_LAMIL Length: 19 December 22, 2002 19:30 Type: P Check: 4317 ..  
 1 ATDYDEEEDD RVKVRDAR  
 !!!A\_ SEQUENCE 1.0  
 ID FIBB\_MAMMU STANDARD: PRT; 21 AA.  
 AC P14475;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Muntiacus muntjak (Muntjak).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
 OC Cervidae; Muntiacinae; Muntiacus.

OX NCBI\_TAXID=9888;  
 RN [1]  
 RP SEQUENCE.  
 RA Gross G.A., Doolittle R.F.;  
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides";  
 RL Arch. Biochem. Biophys. 122:674-684(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 POLIMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 THROMBIN WHICH CLEAVES FIBRINOPETIDES A AND B FROM ALPHA & BETA  
 CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro: IPR002181; Fibrinogen\_C.  
 PROSITE: PS0514; FIBRIN\_AC\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2514 MW; FCFE75188FC0C1627 CRC64;

FIBB\_MUNMU Length: 21 December 22, 2002 19:30 Type: P Check: 6963 ..

1 QHSTDYDEV DDKRLHLDA R

!!!AA SEQUENCE 1.0

ID FIBR\_OPOHE STANDARD; PRT; 21 AA.

AC P14476; 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)

RT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).

GN FGB.

OS Odocoileus hemionus (Mule deer) (Black-tailed deer).  
 OC Eutheria; Chondrocrania; Vertebrata; Euteleostomi;  
 RC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
 RL Arch. Biochem. Biophys. 118:456-467(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 POLIMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 AGGREGATION.

CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 THROMBIN WHICH CLEAVES FIBRINOPETIDES A AND B FROM ALPHA & BETA  
 CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR InterPro: IPR002181; Fibrinogen\_C.

PROSITE: PS0514; FIBRIN\_AC\_DOMAIN; PARTIAL.

KW Blood coagulation; Plasma; Sulfation.

FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.

FT NON\_TER 21 21

FIBB\_RANTA Length: 21 December 22, 2002 19:30 Type: P Check: 6866 ..

1 OHLDAYDEV DDKRLHLDA R

!!!AA SEQUENCE 1.0

ID FIBB\_SYNCA STANDARD; PRT; 21 AA.

AC P14481; 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).

GN FGB.

OS Syncerous caffer (Cape buffalo).  
 OC Eutheria; Chondrocrania; Vertebrata; Euteleostomi;  
 RC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 RL Bovidae; Bovine; Syncerous.

CC NCBI\_TAXID=9970;

RN [1]

RP SEQUENCE.

RA Doolittle R.F., Schubert D., Schwartz S.A.;  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. I."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. II."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. III."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. IV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. V."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. VI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. VII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. VIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. IX."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. X."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XIV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XVI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XVII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XVIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XVIX."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XX."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXIV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXVI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXVII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXVIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXIX."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXX."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXIV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXV."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXVI."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXVII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXVIII."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXIX."  
 CC "Amino acid sequence studies on artiodactyl fibrinopeptides. XXXX."

DR InterPro: IPR002181; Fibrinogen\_C.

PROSITE: PS0514; FIBRIN\_AC\_DOMAIN; PARTIAL.

KW Blood coagulation; Plasma; Sulfation.

FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.

FT NON\_TER 21 21

FIBB\_ODOHE Length: 21 December 22, 2002 19:30 Type: P Check: 6956 ..

1 OHLDAYDEV DDKRLHLDA R

!!!AA SEQUENCE 1.0

ID FIBB\_RANTA STANDARD; PRT; 21 AA.

AC P14479;

SQ SEQUENCE 21 AA: 2353 MN: 09E575AFQ33B863D CRC64:  
 FIBR\_SYNC Length: 21 December 22, 2002 19:30 Type: P Check: 7177 ..  
 1 QFPDYDEGE DDRPKSGLGA R

!-!AA\_SEQUENCE 1.0  
 ID GLU1\_ORENTI STANDARD; PRT: 36 AA.  
 AC P81026;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Glucagon I.  
 OS Oreochromis niloticus (Nile tilapia) (*Tilapia nilotica*).  
 OC Oreochromis metacara; Chondriata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Buteleosteoi; Neoteleosteoi;  
 OC Acanthomorpha; Paracormorpha; Perciformes; Labroidei;  
 OC Cichlidae; Osteichthyes.  
 NCBI\_TAXID=8128;  
 RN 1]  
 RP SQSEQUENCE=  
 RX MEDLINE=95384941; PubMed=7656183;  
 RA Nguyen T.M., Wright J.R. Jr., Nielsen P.F., Conlon J.M.;  
 RT "Characterization of the pancreatic hormones from the Brockmann body  
 of the tilapia: implications for islet xenograft studies";  
 RL Comp. Biochem. Physiol. 11C:33-44(1995).  
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES  
 THE BLOOD SUGAR LEVEL.  
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS  
 CC -!- IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.  
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.  
 DR HSP: P01241; IGCN.  
 DR Interpro-IPO0532; GLUCAGON.  
 DR Pfam: PF00122; GLUCAGON.  
 DR PRINTS: PRO0275; GLUCAGON.  
 DR SMART: SM00070; GLUC.  
 DR PROSITE: PS00260; GLUCAGON; 1.  
 KW Glucagon family; Hormone;  
 SQ SEQUENCE 36 AA: 4252 MW: 54AABFB0254425756 CRC64;  
 GLU1\_ORENTI Length: 36 December 22, 2002 19:30 Type: P Check: 382 ..

!-!AA\_SEQUENCE 1.0  
 ID MY1\_EIF5O STANDARD; PRT: 14 AA.  
 AC P46979;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DR Myoactive tetradecapeptide (EPP).  
 OS Eisentia foetida (Common branding worm) (Common dung-worm).  
 OC Bokayoya; Metzcoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
 OC Lumbricina; Lumbricidae; Eisenia.  
 OC NCBI\_TAXID=6396;  
 RN 1]  
 RP SQSEQUENCE AND SYNTHESIS.  
 RX TISSUE=<gt;  
 RA MEDLINE=95087879; PubMed=85326604;  
 RA Ukena K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,  
 RA Nomoto K.;  
 RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia  
 foetida";  
 RL Peptides 16:995-998(1995).  
 CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT  
 CC MUSCLES.  
 CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.  
 KW Neuropeptide; Amidation.  
 PT MOD\_RES: 14 14  
 SO SEQUENCE 14 AA: 1478 MW: CGC9ABEF941CD91AD CRC64;  
 MY14\_EIF5O Length: 14 December 22, 2002 19:30 Type: P Check: 7600 ..

!-!AA\_SEQUENCE 1.0  
 ID MY14\_PHEVTI STANDARD; PRT: 14 AA.  
 AC P46980;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Myoactive tetradecapeptide (PFP).  
 OS Phereclima victata (Barthwieler); Clitellata; Oligochaeta; Haplotaxida;  
 OC Bokayoya; Metacoel; Annelida; Phereclima.  
 OC Lumbricina; Mebascolecidae; Phereclima.  
 NCBI\_TAXID=46674;  
 RN 1]  
 RP SQSEQUENCE AND SYNTHESIS.  
 RC TISSUE=<gt;  
 RX MEDLINE=95087879; PubMed=85326604;  
 RA Ukena K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,  
 RA Nomoto K.;  
 RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia  
 foetida";  
 RL Peptides 16:995-998(1995).  
 CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT  
 CC MUSCLES.  
 CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.

KW Neuropeptidide; Amidation.  
 FT MOD\_RES 14 14 AMIDATION.  
 SO SEQUENCE 14 AA; 1522 MW; DA40BEB67CCD91AD CRC64;  
 MY14\_PHEVT Length: 14 December 22, 2002 19:30 Type: p Check: 7729 ..

1 GFRDGSDR1 SHGF  
 !!AA\_SEQUENCE 1.0  
 ID T2A\_PARTE STANDARD; PRT; 23 AA.  
 AC Q27173;  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-OCT-2001 (Rel. 40, Last annotation update)  
 DE Trichocyst matrix protein T2-A (Secretory granule protein T2-A) (TMP 2-A) (Fragment).  
 DE (GN T2A).  
 OS paramaecium tetrulaelia.  
 OC Ciliophora; Oligohymenophorea; Peniculida;  
 OC Paramecium;  
 OX NCBI\_TAXID=5888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP PARTIAL SEQUENCE.  
 RC STRAIN=D4-2;  
 RX MEDLINE=96059477; PubMed=7579685;  
 RA Meddedu L., Gautier M.-C., Vaysse L., Houari A., Sperling L.;  
 RT "A large multigene family codes for the polypeptides of the  
 crystaline trichocyst matrix in paramaecium.";  
 RL Mol. Biol. Cell 6:649-659(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP PARTIAL SEQUENCE.  
 RC STRAIN=D4-2;  
 RX MEDLINE=95119139; PubMed=7819344;  
 RA Meddedu L., Gautier M.-C., le Caer J.-P., de Loubresse N., Sperling L.;  
 RT Protein processing and morphogenesis of secretory granules in  
 Paramecium";  
 RL Biochimie 76:329-335(1994).  
 CC -!!- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE  
 TRICHOCYST MATRIX.  
 CC -!!- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY  
 COMPLEX SECRETORY STORAGE GRANULES DOCKED AT THE PLASMA MEMBRANE.  
 CC -!!- SIMILARITY: BELONGS TO THE TMP FAMILY.  
 CC -!!- DATABASE: NAME=Protein Spotlight;  
 CC NOTE="Issue 3 of October 2000."  
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt03.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL: U27510; AAH92610;  
 KW Polyprotein; Structural protein; Multigene family.  
 FT NON\_TER 1  
 FT NON\_TER 23  
 SO SEQUENCE 23 AA; 2706 MW; 184D4BF2759BDB7 CRC64;  
 T2B\_PARTE Length: 23 December 22, 2002 19:30 Type: p Check: 1016 ..

1 DPLDRVSL TDLEDVYVAE QKE  
 !!AA\_SEQUENCE 1.0  
 ID Q9G33 PRELIMINARY; PRT; 17 AA.  
 AC Q9G33;  
 DT 01-MAY-1999 (PREMbrei, 10, Created)  
 DT 01-MAY-1999 (PREMbrei, 10, Last sequence update)  
 DT 01-MAY-1999 (PREMbrei, 10, Last annotation update)  
 DE Subtilisin/chymotrypsin inhibitor (Fragment).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydaceae; Chlamydia.  
 OX NCBI\_TAXID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=L2 434B;  
 RA Wang L., Steenberg S.D., Zheng Y., Larsen S-H;  
 RR "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF087336; RAA04110.1;  
 FT NON\_TER 1  
 FT NON\_TER 17  
 SO SEQUENCE 17 AA; 2043 MW; 4FEE704EE04IE120 CRC64;  
 T2A\_PARTE Length: 23 December 22, 2002 19:30 Type: p Check: 946 ..

! !AA\_SEQUENCE 1.0  
 ID T2B\_PARTE STANDARD; PRT; 23 AA.  
 AC Q27174;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Trichocyst matrix protein T2-B (Secretory granule protein T2-B) (TMP 2-B) (Fragment).  
 DE (GN T2B).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D4-2;  
 RX MEDLINE=96059477; PubMed=7579685;  
 RA Meddedu L., Gautier M.-C., Vaysse L., Houari A., Sperling L.;  
 RT "Protein processing and morphogenesis of secretory granules in  
 Paramaecium";  
 RL Biochimie 76:329-335(1994).  
 CC -!!- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE  
 TRICHOCYST MATRIX.  
 CC -!!- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY  
 COMPLEX SECRETORY STORAGE GRANULES DOCKED AT THE PLASMA MEMBRANE,  
 CC READY TO REPLY TO AN EXOCYTIC STIMULUS.  
 CC -!!- SIMILARITY: BELONGS TO THE TMP FAMILY.  
 CC -!!- DATABASE: NAME=Protein Spotlight;  
 CC NOTE="Issue 3 of October 2000."  
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt03.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL: U27510; AAH92610;  
 KW Polyprotein; Structural protein; Multigene family.  
 FT NON\_TER 1  
 FT NON\_TER 23  
 SO SEQUENCE 23 AA; 2706 MW; 184D4BF2759BDB7 CRC64;

1 NFLGQPFCTD RKNSFRI  
!AA\_SEQUENCE 1.0  
ID Q9KIL6 PRELIMINARY; PRT; 24 AA.  
AC Q9KIL6;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE F01R (Fragment)  
GN F01R.  
OS Streptomyces coelicolor A3(2).  
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
NCBI\_TaxID=100226;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Kormanec J., Sevcikova B., Homerova D.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: AF230491; AAF82064.1; - .  
FT NON\_TER 24 24  
SEQUENCE 24 AA; 2524 MW; FDA1B812ED67583B CRC64;  
Q9KIL6 Length: 24 December 22, 2002 19:30 Type: P Check: 2042 ..  
1 MAADRGTCDE ATAERRALGSR APEF  
!AA\_SEQUENCE 1.0  
ID Q9X9U4 PRELIMINARY; PRT; 25 AA.  
AC Q9X9U4;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE ADP glucose pyrophosphorylase (Fragment).  
GN GIcC.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Bruton C.J.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
MEDLINE=98062210; PubMed=9401038;  
RA Martin M., Schneider D., Bruton C.J., Chater K.F., Hardisson C.;  
RT "A glc Gene Essential Only for the First of Two Spatially Distinct  
RT Phases of Glycogen Synthesis in Streptomyces coelicolor A3(2).";  
RL J. Bacteriol. 179:7784-7789(1997).  
DR EMBL: AJ243801; CAB50742.1; - .  
DR InterPro: IPR001825; NTP\_transferase; 1.  
FT Pram; PF0483; NTP\_transferase; 1.  
FT NON\_TER 25 25  
SEQUENCE 25 AA; 2566 MW; 84B94A728A41D25C CRC64;  
Q9X9U4 Length: 25 December 22, 2002 19:30 Type: P Check: 4039 ..  
1 MGIVLVLAGE GKKMLPLTD RAKPA  
!AA\_SEQUENCE 1.0  
ID Q49137 PRELIMINARY; PRT; 32 AA.  
AC Q49137;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Mxas protein (Fragment).  
GN Mxas.  
OS Methylobacterium extorquens.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Methylobacterium group; Methylobacterium.

OX NCBI\_TaxID=408;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AM1;  
RX MEDLINE=96074311; PubMed=7592474;  
RA Morris C.J., Kim Y.M., Perkins K.E., Lidstrom M.E.;  
RT "Identification and nucleotide sequences of mxaB, mxaC, mxaK, mxaL,  
RT and mxa genes from Methylobacterium extorquens AM1.;"  
RL J. Bacteriol. 177:6825-6831(1995).  
DR EMBL: L41608; AAA85567.1;  
FT NON\_TER 1 1  
SEQUENCE 32 AA; 3700 MW; EB03BB6B3C56ED5 CRC64;

1 ALRRICAPFA RPPFLRADP DAEALSRHLM TT  
!AA\_SEQUENCE 1.0  
ID Q53914 PRELIMINARY; PRT; 9 AA.  
AC Q53914;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Curd protein (Fragment).  
OS Streptomyces cyaneus (Streptomyces curacao).  
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
NCBI\_TaxID=1904;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bergh S.T., Uhlen M.;  
RT "Cloning, analysis and heterologous expression of the polyketides  
RT synthesis genes of Streptomyces curacao.";  
RL Submitted (JUN-1990) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: M33704; AAA26724.1; - .  
FT NON\_TER 1 1  
SEQUENCE 9 AA; 1027 MW; 995BDPDC4140AB1 CRC64;

Q53914 Length: 9 December 22, 2002 19:30 Type: P Check: 3396 ..  
1 ITDERRAAQP  
!AA\_SEQUENCE 1.0  
ID Q50476 PRELIMINARY; PRT; 13 AA.  
AC Q50476;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Catlase.  
GN KATG.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L10373;  
RA Cockerill F.R., Uhl J.R., Temesgen Z., Zhang Y., Stockman L.,  
RA Roberts G.D., Williams D.L., Kline B.C.;  
RT "Rapid identification of a point mutation of the Mycobacterium  
RT tuberculosis catalase-peroxidase (katG) gene associated with isoniazid  
RT resistance.;"  
RL Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: U06263; AAC59976.1; - .  
SEQUENCE 13 AA; 1564 MW; 2F39A45FEE94777 CRC64;  
Q50476 Length: 13 December 22, 2002 19:30 Type: P Check: 7152 ..  
1 MPEOHPPITD RSR  
!AA\_SEQUENCE 1.0  
ID 005422 PRELIMINARY; PRT; 38 AA.  
AC 005422;

DT 01-JUL-1997 (TREMBrel. 04, Created)  
 DT 01-JUL-1997 (TREMBrel. 04, Last sequence update)  
 DT 01-NOV-1998 (TREMBrel. 08, Last annotation update)  
 DE Phytoene dehydrogenase (Fragment).  
 GN CPTI.  
 OS Mycobacterium marinum.  
 OC Bacteria; Firmicutes; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TaxID=1781;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M;  
 RX MEDLINE=97440138; PubMed=9294446;  
 RA Ramakrishnan L., Tran H.T., Federspiel N.A., Falkow S.;  
 RT "A CRTB homolog essential for Photochromogenicity in *Mycobacterium*  
 marinum: isolation, characterization, and gene disruption via  
 homologous recombination."  
 RL J. Bacteriol. 179:5862-5868(1997).  
 DR EMBL; U92075; AMB71427.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 38 AA; 3986 MW; 6E46332707CCDCAB CRC64;  
 005422 Length: 38 December 22, 2002 19:30 Type: P Check: 6757 ..  
 1 VPGVGVPPTL ISGRLLAADDRTGNTTRSRH LDLKAQLS  
 ! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 19 AA.  
 ID Q93A12; PRT; 22 AA.  
 AC 093A12; PRT; 22 AA.  
 DT 01-DEC-2001 (TREMBrel. 19, Created)  
 DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE Putative ribitol/glucose oxidoreductase (Fragment).  
 GN SDR2A.  
 OS Thiobacillus ferrooxidans.  
 OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
 OC NCBI\_TaxID=920;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC33020;  
 RA Bruscella P., Levican G., Ratouchniak J.J., Holmes D., Bonnefoy V.;  
 RT "A second operon encoding a bcl complex in *Acidithiobacillus*  
 ferrooxidans." Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ311888; CAC44744.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 22 AA; 2464 MW; 3725F8E43BBA75C2 CRC64;  
 093A12 Length: 22 December 22, 2002 19:30 Type: P Check: 9419 ..  
 1 AGNFNFNPDD DRTEKDLATS TV  
 ! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 12 AA.  
 ID Q93A08; PRT; 12 AA.  
 AC 093A08; PRT; 12 AA.  
 DT 01-DEC-2001 (TREMBrel. 19, Created)  
 DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE RSEB protein (Fragment).  
 GN RSEB.  
 OS Thiobacillus ferrooxidans.  
 OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
 OC NCBI\_TaxID=920;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC33020;  
 RA Levican G., Bruscella P., Guacunano M., Inostroza C., Jedlicki E.,  
 RA Bonnefoy V., Holmes D.S.;  
 RT "Characterization of the pet and res operons of *Acidithiobacillus*  
 ferrooxidans." Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ413194; CAC88360.1; -.  
 FT NON\_TER 1 1

---

SQ SEQUENCE 12 AA; 1405 MW; 886AB7DFE13240A CRC64;  
 ID Q93A08 Length: 12 December 22, 2002 19:30 Type: P Check: 5753 ..  
 1 QSQDRKENN DG  
 ! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 19 AA.  
 ID Q03977; PRT; 19 AA.  
 AC 003977; PRT; 19 AA.  
 DT 01-NOV-1996 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBrel. 08, Last annotation update)  
 DE Alpha-acceptor polypeptide M15 (Fragment).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D15 ALPHA;  
 RX MEDLINE=93083990; PubMed=1339377;  
 RA Prentki P.C.;  
 RT "Nucleotide sequence of the classical lacZ deletion delta M15.";  
 RL Gene 122:231-232(1992).  
 DR X58252; CAA41206.1; -.  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 2092 MW; 8324315E003AA053 CRC64;  
 003977 Length: 19 December 22, 2002 19:30 Type: P Check: 4923 ..  
 1 MTMTDSLAV VARTDRPSQ  
 ! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 36 AA.  
 ID Q99094; PRT; 36 AA.  
 AC 099094; PRT; 36 AA.  
 DT 01-NOV-1996 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBrel. 02, Last annotation update)  
 DE MKA protein (Fragment).  
 OS Salmonella typhimurium.  
 OG Plasmid PYA426.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OC NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Guillig P.A., Chiodo V.A.;  
 RL Infect. Immun. 58:2651-2658(1991).  
 DR EMBL; M64295; AAA27272.1; -.  
 KW Plasmid.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 36 AA; 4121 MW; C3D43D1C622EBAB7 CRC64;  
 099094 Length: 36 December 22, 2002 19:30 Type: P Check: 1642 ..  
 1 GSQDFASQLS KURLSDDRTA DTNRRIKRIN MRVLNS  
 ! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 20 AA.  
 ID Q9N25; PRT; 20 AA.  
 AC 09N25; PRT; 20 AA.  
 DT 01-OCT-2000 (TREMBrel. 15, Created)  
 DT 01-OCT-2000 (TREMBrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
 DE ATP7B (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Cetartiodactyla; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wu Z.Y., Wang N., Murong S.X., Lin M.T., Fang L.;  
 RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF220215; AAF67661.1; -.

DR InterPro; IPR001757; ATPase\_E1-E2.  
 DR Pfam; PF00122; E1-E2\_ATPase; 1.  
 FT NON\_TER 1 1  
 FT VARIANT 19 19 V -> I.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2282 MW; 00268E2C1E0692E3 CRC64;  
 09NZ25 Length: 20 December 22, 2002 19:30 Type: P Check: 6136 ..

1 APIQOLADRF SGYFVPFIVI

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRTR; 40 AA.

ID Q15244

AC 015244

DT 01-NOV-1995 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMBrel. 15, Last annotation update)

DE Phosphoribosylpyrophosphate synthetase isoform I (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. [1]

NCBI\_TaxID=9606;

RN RC

RN TISSUE=BLOOD;

RN Ishizuka T.;

RN Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases. [2]

RN RPP

RN SEQUENCE FROM N.A.

RN TISSUE=BLOOD;

RN Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shimada H., Nagatake N., Tatsubana M.;

RN "Promoter regions of the human X-linked housekeeping genes PRPS1 and PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II isoforms."; Blochim. Biophys. Acta 1130:139-148(1992).

RN EMBL; D28133; BRA05675.1; -. HSSP; P14193; IDKU.

RN FT NON\_TER 40

SO Q15244 Length: 40 December 22, 2002 19:30 Type: P Check: 2868 ..

1 MPNIKIFSGS SHQDLSQRTA DRGLGELGKV VTKKFSNQET

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRTR; 40 AA.

ID Q15245

AC 015245

DT 01-NOV-1995 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)

DE Phosphoribosylpyrophosphate synthetase isoform II (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. [1]

NCBI\_TaxID=9605;

RN RPP

RN SEQUENCE FROM N.A.

RN TISSUE=BLOOD;

RN Ishizuka T.;

RN Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases. [2]

RN RPP

RN SEQUENCE FROM N.A.

RN TISSUE=BLOOD;

RN MEDLINE=92223087; PubMed=1314091;

RN Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shimada H., Nagatake N., Tatsubana M.;

RN "Promoter regions of the human X-linked housekeeping genes PRPS1 and PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II isoforms."; Blochim. Biophys. Acta 1130:139-148(1992).

RN EMBL; D28134; BA05676.1; -. [1]

FT NON\_TER 40 SQ SEQUENCE 40 AA; 4431 MW; CBC2B31685083EFF CRC64;

Q15245 Length: 40 December 22, 2002 19:30 Type: P Check: 3314 ..

1 MNIVILFSGS SHQDLSQRTA DRGLGELGKV VTKKFSNQET

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRTR; 34 AA.

ID Q9BQVB

AC 09BQVB

DT 01-JUN-2001 (TREMBrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBrel. 21, Last annotation update)

DE L-isospartyl/D-aspartyl methyltransferase (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. [1]

NCBI\_TaxID=9606;

RN RPP

RN SEQUENCE FROM N.A.

RN Tissue W., Clarke S.; PRMT; 1.

RN MEDLINE=91354847; PubMed=8074695;

RN Tsai W., Clarke S.; PRMT; 1.

RN "Amino acid polymorphisms of the human L-isospartyl/D-aspartyl methyltransferase involved in protein repair"; Biochem. Biophys. Res. Commun. 203:491-497(1994).

RN DR EMBL; S73902; AAC60639; 2; -.

RN DR InterPro; IPR000682; Pint.

RN PRMT; 1.

RN KW Methylytransferase; Transferase.

FT NON\_TER 34 SQ SEQUENCE 34 AA; 3873 MW; EEB18E55F49BF377 CRC64;

Q9BQVB Length: 34 December 22, 2002 19:30 Type: P Check: 5960 ..

1 NGIIKTDKFV EMVLMATRSH YAKCPYMDS POSI

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRTR; 15 AA.

ID Q9CQCI

AC Q9CQCI

DT 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBrel. 17, Last annotation update)

DE Tropomyosin (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. [1]

NCBI\_TaxID=9606;

RN RPP

RN SEQUENCE

RX MEDLINE=93195352; PubMed=8450225;

RX Das K.M., Dasgupta A., Mandal A., Geng X.;  
 RT "Autoimmunity to cytoskeletal protein tropomyosin. A clue to the pathogenetic mechanism for ulcerative colitis."; J. Immunol. 150:2487-2493(1993).

RN DR InterPro; IPR000533; Tropomyosin.

RN Pfam; PF00261; Tropomyosin.

SO SEQUENCE 15 AA; 1802 MW; 7A0993CA5A54254C CRC64;

Q9CQCI Length: 15 December 22, 2002 19:30 Type: P Check: 8961 ..

1 HIAEDADKRY EEVVR

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRTR; 28 AA.

ID Q9FR68

AC Q9FR68

DT 01-DEC-2001 (TREMBrel. 19, Created)  
 DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)

DE UnKnown (Protein for IMAGE:563468) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. [1]

OX NCBI\_TaxID=3606;  
 RN [1]  
 RP PRELIMINARY;  
 SEQUENCE FROM N.A.  
 TISSUE="KIDNEY";  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BCO11565; RAHL1565.1;  
 FT NON\_TER 1  
 SEQUENCE 28 AA; 2841 MW; 5799D138245D3951 CRC64;  
 SQ [1]  
 096F68 Length: 28 December 2, 2002 19:30 Type: P Check: 588 ..  
 1 VSQPGSCRH ADRVGHVGQR AGAGVRPE  
 RN [1]  
 RP PRELIMINARY;  
 ID 09NBB1 PRT; 26 AA.  
 AC 09NBB1:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Stretchin-MICK (Fragment)  
 DR STRN-MICK OR CG8304 OR CG18255.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephdroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7277;  
 RN [1]  
 RPP SEQUENCE FROM N.A.  
 RA Champagne M.B.; Edwards K.A.; Erickson H.P.; Kiehart D.P.; Kiehart D.P.; Leucosolenida;  
 RT "Drosophila stretchin-MICK Is a Novel Member of the Titin/Myosin Light  
 Chain Kinase Family";  
 RL J. Mol. Biol. 0-0-01(2000).  
 DR EMBL: AR25709; AAF90127.1; -.  
 DR FlyBase: FBgn0013988; STRN-MICK.  
 FT NON\_TER 1  
 SQ SEQUENCE 26 AA; 3004 MW; 2F2CB3A55B7F033 CRC64;  
 09NBB1 Length: 26 December 22, 2002 19:30 Type: P Check: 6662 ..  
 1 VAWNNNGFTD RIFVVTIKI PKKKEE  
 !\_!AA\_SEQUENCE 1.0  
 ID 09G045 PRELIMINARY; PRT; 27 AA.  
 AC 09G045;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Homeobox protein Sycox2 (Fragment).  
 GN SYCOX2.  
 OS Sycon raphanus.  
 OC Syconidae; Metazoa; Porifera; Calcarea; Calcaronea; Leucosolenida;  
 OC NCBI\_TaxID=56443;  
 RN [1]  
 RPP SEQUENCE FROM N.A.  
 RX MEDLINE=20476355; PubMed=11020308;  
 RA Manuel M.; Le Parco Y.;  
 RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon  
 raphanus";  
 RL Mol. Phylogenet. Evol. 17:97-107(2000).  
 DR EMBL: AF137141; AAG28511.1; -.  
 DR InterPro: IPR001336; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 27 AA; 3014 MW; CA58E82803F29E3B7 CRC64;  
 09GU44 Length: 27 December 22, 2002 19:30 Type: P Check: 9332 ..  
 1 FTKSKLAPPT RGALAARLGL TDQVKV  
 !\_!AA\_SEQUENCE 1.0  
 ID 09X226 PRELIMINARY; PRT; 23 AA.  
 AC 09X226;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE Calmodulin (Fragment);  
 GN CAM.  
 OS Littorina saxatilis.  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Mesogastropoda; Littorioidea; Littorinidae; Littorina.  
 NCBI\_TaxID=31220;  
 RN [1]  
 RPP SEQUENCE FROM N.A.  
 RA Wilding C.S.; Graeme J.; Mill P.J.;  
 RT "Molecular characterisation of calmodulin intron variation in  
 Littorina (Gastropoda: Prosobranchia) species.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: AJ133335; CAB44224.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 23 AA; 2815 MW; 5762B9C4E9886A06 CRC64;  
 09X226 Length: 23 December 22, 2002 19:30 Type: P Check: 1592 ..  
 1 LDQDMNEVDA DRQDRDLPR VPY  
 !\_!AA\_SEQUENCE 1.0  
 ID 094631 PRELIMINARY; PRT; 25 AA.  
 AC 094631;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PmBbox1 (Fragment);  
 OS Polyantracarpa misakiensis  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiaceae;  
 OC Stolidobranchia; Styelidae; Polyantracarpa.  
 NCBI\_TaxID=7723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USHIMADO;  
 RA Fujiiwara S.; Kawamura K.;  
 RT "Cloning of homeobox-containing genes from the budding ascidian  
 Polyandrocarpa misakiensis.";

RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: DE8274; BA113570.1; -.  
 DR InterPro: IPR001355; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 FT NON\_TER 25 25 ID 1  
 SEQUENCE 25 AA; 3074 MW; D7CBE263F4B30A51 CRC64;  
 094681 Length: 25 December 22, 2002 19:30 Type: P Check: 5282 ..

1 HFNQYLCRER RQEVARAKVNL TDROVKI

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID 094681 PRELIMINARY; PRT; 27 AA.  
 AC 094681  
 DT 01-FEB-1997 (TREMBrel. 02, Created)  
 DT 01-FEB-1997 (TREMBrel. 02, Last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
 DE Ovxl ortholog homeobox (Fragment).  
 GN CTS-OVXL  
 OS Ctenodrilus serratus.  
 OS Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;  
 OC Terebellida; Ctenodrilidae; Ctenodrilus.  
 OX NCBI\_TaxID=40316;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9436262; PubMed=7915607;  
 RA DICK M.H., BUSS L.W.;  
 RT "A PCR-based survey of homeobox genes in Ctenodrilus serratus  
 (Annelida: Polychaeta)." ;  
 RL Mol. Phylogenet. Evol. 3:146-158(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9436262; PubMed=7915607;  
 RA DICK M.H., BUSS L.W.;  
 RT Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: U26634; AAC46856.1; -.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SEQUENCE 27 AA; 2879 MW; 4CCE22303AA0102 CRC64;  
 Q23749 Length: 27 December 22, 2002 19:30 Type: P Check: 8475 ..

1 POBILVTDG TIAHRASPET DRGSGEN

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID 025482 PRELIMINARY; PRT; 27 AA.  
 AC 025482  
 DT 01-NOV-1995 (TREMBrel. 01, Created)  
 DT 01-MAY-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE Anthox. Ms (Fragment).  
 GN ANTHOXI.MS.  
 OS Metridium senile (Brown sea anemone) (Frilled sea anemone).  
 OC Nyanthae; Metazoa; Anthozoa; Zoantharia; Actiniaria;  
 OC NCBI\_TaxID=6116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97435515; PubMed=9290214;  
 RA Finnerity J.R., Mantindale M.Q.;  
 RT "Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of  
 Nematocilla vectensis and Metridium senile." ;  
 RL Biol. Bull. 193:65-76(1997).  
 DR EMBL: U42727; AAC46626.1; -.  
 DR InterPro: IPR001355; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINS; PR00024; HOMEBOX.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SEQUENCE 27 AA; 3314 MW; E6BD0191EE3905CF CRC64;  
 Q25482 Length: 27 December 22, 2002 19:30 Type: P Check: 9269 ..

1 HNFELFLAKER RESEMATQNL TDROVKI

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID 094558 PRELIMINARY; PRT; 27 AA.  
 AC 094558  
 DT 01-FEB-1997 (TREMBrel. 02, Created)  
 DT 01-FEB-1997 (TREMBrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE CHV-HBL protein (Fragment).  
 GN CHV-HBL  
 OS Chaetopterus variopedatus.  
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;  
 OC Spiorida; Chaetopteridae; Chaetopterus.  
 OX NCBI\_TaxID=34590;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA IRVINE S.M., WARMERD A.J., HUNTER J.D., MARTINALE M.Q.;  
 RT "A survey of homeobox genes in Chaetopterus variopedatus and analysis  
 of polychaete homeodomains";  
 RT Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: U068283; AAC16992.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SEQUENCE 27 AA; 3351 MW; 74E5C77ED08CDEE4 CRC64;  
 SQ 094458 Length: 27 December 22, 2002 19:30 Type: P Check: 8578 ..

1 NOKKIEKKD RDRISNEIGL DDQTKY

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID 096592 PRELIMINARY; PRT; 27 AA.  
 AC 096592  
 DT 01-MAY-1999 (TREMBrel. 10, Created)  
 DT 01-MAY-1999 (TREMBrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE Caudal (Fragment).  
 GN CAD.  
 OS Sacculina carcinii.  
 OC Rhirocephala; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;  
 OX NCBI\_TaxID=51650;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98334561; PubMed=9667986;  
 RA Moucheil-Viehn E., Rigolot C., Gilbert J.M., Deutsch J.S.;  
 RT "Molecules and the body plan: the Hox genes of Cirripedes  
 (Crustacea)." ;  
 RL Mol. Phylogenet. Evol. 9:382-398(1998).  
 DR EMBL: U73471; AAC00342.1; -.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00045; homeobox; 1.  
 DR PRINS; PR00024; HOMEBOX.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SEQUENCE 27 AA; 3289 MW; 1CDC2EACAO4E00F CRC64;  
 SQ 096892 Length: 27 December 22, 2002 19:30 Type: P Check: 9212 ..

1 RFNNYITIKR KLELSRIGL TDROVKI

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID P91901 PRELIMINARY; PRT; 27 AA.  
 AC P91901  
 DT 01-MAY-1997 (TREMBrel. 03, Created)  
 DT 01-MAY-1997 (TREMBrel. 03, Last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
 DE Homeobox protein BHOx26 (Fragment).  
 OS Beroe ovata.  
 OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroidea; Beroidae; Beroe.  
 OX NCBI\_TaxID=10201;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Witchell H.J., Meach R.W.;  
 RT "BHOx26 - Beroe Homeobox;"  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DDJB databases.  
 DR EMBL; U83381; AAB9471.1;  
 DR InterPro: IPR001556; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS: PR0024; HOMEBOX.  
 DR PROBOM: PD000010; Homeobox; 1.  
 FT NON\_TER 1 1  
 FT 27 27  
 SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

P91901 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1. LFNNYLTTRER RLEISKINSNL TDQVKI

! !AA\_SEQUENCE 1.0  
 ID 002602 PRELIMINARY; PRT; 27 AA.  
 AC 002602; 002335;  
 DT 01-JUL-1997 (TREMBREL. 04, Created)  
 DT 01-JUL-1997 (TREMBREL. 04, Last sequence update)  
 DT 01-MAR-2002 (TREMBREL. 20, Last annotation update)  
 DE Homeobox protein BHox35 (fragment).  
 OS Beroe ovata.  
 OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroida; Beroidae; Beroe.  
 OC NCBI\_TAXID=10201;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Witchell H.J., Meach R.W.;  
 RA "BHox35 - Beroe homeobox;"  
 RA Submitted (FEB-1997) to the EMBL/GenBank/DDJB databases.  
 DR EMBL; U83383; AAB9473.1; -.  
 DR InterPro: IPR001556; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS: PR0024; HOMEBOX.  
 DR PROBOM: PD000010; Homeobox; 1.  
 FT NON\_TER 1 1  
 FT 27 27  
 SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

002602 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..

1. LFNNYLTTRER RLEISRGVNL TDQVKI

! !AA\_SEQUENCE 1.0  
 ID 018615 PRELIMINARY; PRT; 34 AA.  
 AC 018615;  
 DT 01-JAN-1998 (TREMBREL. 05, Created)  
 DT 01-JUN-1998 (TREMBREL. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMBREL. 21, Last annotation update)  
 DR FTZ\_F1 (Fragment).  
 OS Artemia salina (Brine shrimp).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;  
 OC Artemioidea; Artemia.  
 OC NCBI\_TAXID=83549;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RM MEDLINE=9733099; PubMed=9192646;  
 RX Escriba H., Safi R., Hanni C., Langlois M.C., Saumitou-Laprade P.,  
 RA Stéphain D., Capron A., Pierce R., Lauden V.,  
 RT "Ligand binding was acquired during evolution of nuclear receptors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:6003-6808(1997);  
 DR -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY);  
 DR -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL; U93418; AAC83395.1; -.  
 DR InterPro: IPR001528; Znf\_C4steroid.  
 DR Pfam: PF00105; Zf-C4; 1.  
 DR PROBOM: PD000035; Znf\_C4steroid; 1.  
 DR SMART; SM00399; Znf\_C4; 1.  
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger. 1 1  
 FT NON\_TER 34 34

! !AA\_SEQUENCE 1.0  
 ID 09V973 PRELIMINARY; PRT; 21 AA.  
 AC 09V973;  
 DT 01-MAY-2000 (TREMBREL. 13, Created)  
 DR 01-MAY-2000 (TREMBREL. 13, Last sequence update)  
 DR 01-MAY-2000 (TREMBREL. 13, Last annotation update)  
 DR 01-MAY-2000 (TREMBREL. 13, Last annotation update)  
 DE CG12485 protein.  
 DE CG12485.

SQ SEQUENCE 34 AA; 4132 MW; 1AE9E6A0D63D4AEF CRC64;

018615 Length: 34 December 22, 2002 19:30 Type: P Check: 5276 ..

1. FKRVQNKV YTCMADRSCH IDKSQRKCP PCRF

! !AA\_SEQUENCE 1.0  
 ID 09V973 PRELIMINARY; PRT; 17 AA.  
 AC 09V973;  
 DT 01-MAY-2000 (TREMBREL. 13, Created)  
 DR 01-MAY-2000 (TREMBREL. 13, Last sequence update)  
 DR 01-MAY-2000 (TREMBREL. 13, Last annotation update)  
 DE CG12485.

RX Eukaryota; Metazoa; Anthroponida; Tracheata; Hexapoda; Insecta;  
 OC Drosophila melanogaster (fruit fly).  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydioidea; Dracophilidae; drosophila.  
 OC NCBI\_TAXID=7225;  
 RN [1]  
 RP STRAIN=BERKELEY;

RA MEDLINE=2016006; PubMed=10731132;  
 RA Adams M.D., Celmin S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.M., Yandell M.D., Zheng Q., Chen L.X.,  
 RA Wan K.H., Doyle C., Baxter E.G., Heitz G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbavani A., An H.-J., Andrews Pfrankoch C., Baldwin C., Baldwin C.,  
 RA Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beesley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Borikova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,  
 RA Burts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cowley S., Dahake R., Davenport L.B., Davies P.,  
 RA de Pablo J., Deicher A., Deng Z., Dews A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostier C., Gabrilian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li J., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.R., Nixon K., Nusslein-D.R., Pacelli J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Sider-Kiamos I., Simpson M., Skupski M., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svärksas R., Tector C., Turner E., Venter J., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of Drosophila melanogaster.";  
 RA Science 287:218-219(2000);  
 DR EMBL; A003791; AAF57423.1; -.  
 DR FlyBase; FBgn0034547; CG12485;  
 DR Sequence 17 AA; 1907 MW; 47D598D29F4860DE CRC64;

09V973 Length: 17 December 22, 2002 19:30 Type: P Check: 1645 ..

1. MELSSDRGP HHAQQGR

01-MAY-2000 (TREMBrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBrel. 14, Last annotation update)  
 DE Cysteine proteinase (Fragment).  
 OS Bombyx mori (Silk moth)  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Diptaria;  
 OC Bombycoidea; Bombycidae; Bombyx.  
 NCBI\_TaxId=7091;

[1]  
 !AA\_SEQUENCE 1.0  
 RX MEDLINE=94086476; PubMed=8262908;  
 RA Takahashi S.Y., Yamamoto Y., Shionoya Y., Kageyama T.;  
 RT "Cysteine proteinase from the eggs of the silkworm, *Bombyx mori*:  
 identification of a latent enzyme and characterization of activation  
 and proteolytic processing in vivo and in vitro.";  
 RL J. Biochem. 114:267(1993);  
 SQ SEQUENCE 21 AA; 2379 MW; 9232784B696AE4B1 CRC64;

Q9TWRL Length: 21 December 22, 2002 19:30 Type: P Check: 8106 ..

1 XPEQQVDDRKH GAVIDDKXXQ X

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 26 AA.  
 ID 062575; PRT; 26 AA.  
 AC 062575;  
 DT 01-AUG-1998 (TREMBrel. 07, Created)  
 DT 01-AUG-1998 (TREMBrel. 07, Last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
 DE Hox protein MYX4 (Fragment).  
 GN HOX OR MYX4.  
 OS Tetracapsula bryozoides.  
 OC Eukaryota; Metazoa; Myxozoa; Myxozoa incertae sedis; Tetracapsula.  
 NCBI\_TaxId=75694;

RN [1]  
 !AA\_SEQUENCE FROM N.A.  
 RX MEDLINE=98196655; PubMed=9537319;  
 RA Anderson C.L., Cannon E.U., Okamura B.;  
 RT "A triploblast origin for Myxozoa?";  
 RL Natur 392:346-347(1998)  
 DR InterPro: IPR001356;  
 DR Pfam: PF000010; Homeobox; 1;  
 DR Prodrom; PD000010; Homeobox; 1;  
 KW DNA-binding; Homeobox; Nuclear protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 26 26  
 SQ SEQUENCE 26 AA; 3212 MW; E38651AD1657BD2C CRC64;

Q92575 Length: 26 December 22, 2002 19:30 Type: P Check: 7242 ..

1 LFNHMLRSER RLTSKSTDL TDQVK

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID 061676; PRT; 27 AA.  
 AC 061676;  
 DT 01-NOV-1996 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
 DE HOX11-D125 protein (D125) (Fragment).  
 GN HOX11-D125 OR D125.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyoidea; Drosophilidae; Drosophila.  
 NCBI\_TaxId=7227;

RN [1]  
 !AA\_SEQUENCE FROM N.A.  
 RX MEDLINE=93281533; PubMed=8099440;  
 RA Dear T.N., Sanchez-Garcia I., Rabbits T.H.;  
 RT "The HOXA1 gene encodes a DNA-binding nuclear transcription factor  
 belonging to a distinct family of homeobox genes.";  
 RL Prog. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).  
 EMBL; D08618; AAA28612.1; -.

RP SEQUENCE FROM N.A.  
 RX MEDLINE=93281533; PubMed=8099440;  
 RA Dear T.N., Sanchez-Garcia I., Rabbits T.H.;  
 RT "The HOXA1 gene encodes a DNA-binding nuclear transcription factor  
 belonging to a distinct family of homeobox genes.";  
 RL Prog. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).  
 EMBL; D08618; AAA28612.1; -.

DR FlyBase; FBgn0010394; HOX11-D125.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR Prodrom; PD000010; Homeobox; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 2937 MW; CABRC2F676F2B3A3 CRC64;

Q61676 Length: 27 December 22, 2002 19:30 Type: P Check: 8698 ..

1 LYQKYLSPAD RDETAASIGL SNAQVIT

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 29 AA.  
 ID Q9MZW8; PRT; 29 AA.  
 AC 09MZW8;  
 DR Goodman M.; Sampalo I., Schneider M.P., Page S.L.,  
 RT "The place of Callimico goedeli in the Callitrichine phylogenetic  
 tree: evidence from von Willebrand factor gene intron II sequences.";  
 RL Mol. Phylogenet. Evol. 13:392-404(1999).  
 DR EMBL; AF092833; AAF7501.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 29 29  
 SQ SEQUENCE 29 AA; 3227 MW; 60F1190C8F227D03 CRC64;

Q9MNW8 Length: 29 December 22, 2002 19:30 Type: P Check: 2931 ..

1 GRDCQDHFS IVIETVQCAD DRDAVCTRS

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 23 AA.  
 ID Q23934; PRT; 23 AA.  
 AC Q23934;  
 DT 01-NOV-1996 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE Growth hormone receptor (Fragment).  
 GN GHR.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 NCBI\_TaxId=9615;

RN [1]  
 !AA\_SEQUENCE FROM N.A.  
 RX MEDLINE=9704323; PubMed=8894053;  
 RA Venta P.J., Brouillette J.A., Yuzbasiyian-Gurkan V., Brewer G.J.;  
 RT "Gene-specific universal mammalian sequence-tagged sites: application  
 to the canine genome.";  
 RL Biocem. Genet. 33:321-341(1996).  
 DR EMBL; L177673; AAA97423.1; -.  
 KW Receptor.  
 FT NON\_TER 1 1  
 FT NON\_TER 23 23  
 SQ SEQUENCE 23 AA; 2620 MW; 85442F0EC46B5D08 CRC64;

Q929394 Length: 23 December 22, 2002 19:30 Type: P Check: 807 ..

1 DDLDIBKTEGS DTDRLNSDH EKS

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 24 AA.  
 ID Q9TR57; PRT; 24 AA.  
 AC Q9TR57;  
 DR 01-MAY-2000 (TREMBrel. 13, Created)

01-MAY-2000 (TREMBrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBrel. 21, Last annotation update)  
 DE Vitronectin (Fragment).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 RN [1]  
 RP  
 SEQUENCE:  
 RX MEDLINE=92207982; PubMed=1372829;  
 RA Nakashima N., Miyazaki I., Isikawa M., Yatohgo T., Ogawa H.,  
 RA Uchibori H., Matsumoto I., Seno N., Hayashi M.;  
 RT "vitronectin diversity in evolution but uniformity in ligand binding  
 and size of the core polypeptide.";  
 RL Blochin. Biophys. Acta 1120:1-10(1992).  
 DR InterPro; IPR000886; ER-target.  
 DR PROSITE; PS0014;  
 FT NON\_TER 1 1  
 SQ SEQUENCE 24 AA; 2745 MW; 94F0054986FE1215 CRC64;  
 Q9TR87 Length: 24 December 22, 2002 19:30 Type: P Check: 2610 ..  
 1 AEQESXKGRVT EGFNADRKKQ QDEL  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 15 AA.  
 ID Q9TQ09  
 AC 09TQ09;  
 DT 01-MAY-2000 (TREMBrel. 13, created)  
 DT 01-MAY-2000 (TREMBrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBrel. 14, last annotation update)  
 DE Glutamate dehydrogenase Isoform 1 (EC 1.4.1.2) (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 RN NCBI\_TaxID=9913;  
 RN [1]  
 RP  
 SEQUENCE:  
 RX MEDLINE=96061967; PubMed=7588764;  
 RA Cho S.W., Lee J., Choi S.Y.;  
 RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine  
 brain.";  
 RL Eur. J. Biochem. 233: 340-346 (1995).  
 RN [2]  
 RP  
 SEQUENCE:  
 RX MEDLINE=96043916; PubMed=7581004;  
 RA Lee J., Kim S.W., Cho S.W.;  
 RT "A novel glutamate dehydrogenase from bovine brain: purification and  
 characterization";  
 RL Blochem. Mol. Biol. Int. 36: 1087-1096 (1995).  
 SQ SEQUENCE 15 AA; 1754 MW; 65F7CD91023AEBA CRC64;  
 Q9TRQ9 Length: 15 December 22, 2002 19:30 Type: P Check: 8644 ..  
 1 EEAADREDD PNFFK  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 15 AA.  
 ID Q9TR40  
 AC Q9TR40;  
 DT 01-MAY-2000 (TREMBrel. 13, created)  
 DT 01-MAY-2000 (TREMBrel. 13, last sequence update)  
 DE Glutamate dehydrogenase Isoform II (EC 1.4.1.2) (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 RN NCBI\_TaxID=9913;  
 RN [1]  
 RP  
 SEQUENCE:  
 RX MEDLINE=96061967; PubMed=7588764;  
 RA Cho S.W., Lee J., Choi S.Y.;  
 RN  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 33 AA.  
 ID Q9TR2;  
 AC 09TR2;  
 DT 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBrel. 21, Last annotation update)  
 DE VAMP/synaptobrevin-2 (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 Bovidae; Bovinae; Bos.  
 RN NCBI\_TaxID=9913;  
 RN [1]  
 RP  
 SEQUENCE:  
 RX MEDLINE=93374072; PubMed=8355494;  
 RA Horikawa H.P., Saisu H., Ishizuka T., Sekine Y., Tsugita A., Odani S.,  
 RA Abe T.;  
 RL FEBS Lett. 330: 236-240 (1993).  
 DR InterPro; IPR001388; Synaptobrevin.  
 DR Pfam; PF00957; synaptobrevin; 1.  
 DR ProDom; PD001229; Synaptobrevin; 1.  
 DR PROSITE; PS00117; SYNATOBREVIN; 1.  
 SQ SEQUENCE 33 AA; 3672 MW; 584386688353BC55 CRC64;  
 Q9TRF2 Length: 33 December 22, 2002 19:30 Type: P Check: 1454 ..  
 1 XVNVRKLYER DOKLSEUDDR ADALQAGASO FET  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 22 AA.  
 ID Q37112  
 AC 037112;  
 DT 01-NOV-1995 (TREMBrel. 01, Created)  
 DT 01-NOV-1995 (TREMBrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBrel. 08, Last annotation update)  
 RL ORF22.  
 OS Pinus thunbergii (Green Pine) (Japanese black Pine).  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OC NCBI\_TaxID=3350;  
 RN [1]  
 RP  
 SEQUENCE FROM N.A.  
 RX MEDLINE=92212283; PubMed=1557027;  
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,  
 RA Wakasugi T., Sugiyama M.;  
 RT "Chloroplast DNA of black pine retains a residual inverted repeat  
 lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and  
 trnB and the absence of rps16.";  
 RL Mol. Gen. Genet. 232: 206-214 (1992).  
 RN  
 SEQUENCE FROM N.A.  
 RX MEDLINE=95094112; PubMed=8001170;  
 RA Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugiyama M.;  
 RT "A new gene encoding tRNA(Pro) (GGG) is present in the chloroplast  
 genome of black pine: a compilation of 32 tRNA genes from black pine  
 chloroplasts.";  
 RT Curr. Genet. 26: 153-158 (1994).  
 RN  
 SEQUENCE FROM N.A.  
 RX MEDLINE=95024047; PubMed=7937893;  
 RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,  
 RA Sugiyama M.;  
 RT "Loss of all ndh genes as determined by sequencing the entire  
 chloroplast genome of the black pine Pinus thunbergii.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:3794-3798(1994).  
RN [4]  
RN  
RP SEQUENCE FROM N.A.  
RA Suglira M.;  
RL Submitted (SEP-1994) to the EMBL/GenBank/DDAJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94138245; PubMed=8305874;  
RA Li J., Goldschmidt-Clermont M., Timko M.;  
RT "Chloroplast-encoded clib is required for light-independent protochlorophyllide reductase activity in *Chlamydomonas reinhardtii*."  
RT Plant Cell 5:1817-1829(1993).  
DR EMBL: D17510; BAA0307..;  
DR EMBL: D11467; BAA02023.1; -.  
KW Chloroplast.  
SQ SEQUENCE 22 AA; 2606 MW; 0A95CB0443BCAEC5 CRC64;  
Q37112 Length: 22 December 22, 2002 19:30 Type: P Check: 9119 ..

1 MELITDBRSI ECGIVLKKIE SI

! ! AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 22 AA.

ID Q957T4  
AC Q957T4;  
DT 01-DEC-2001 (TREMBrel. 19, Created)  
DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)

DE Ribosomal protein S12 (Fragment).  
GN RPS12.  
OS Pinus sylvestris (Scots pine).  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus;  
OX NCBI\_TaxID=3349;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wachowiak W., Liepelt S., Prus-Glowacki W.;  
RT "Sequence analysis of chosen regions of the mitochondrial genome of  
RL Pinus mugo and P. sylvestris.";  
DR EMBL: AE359577; ARK3919.1; -.  
KW Mitochondrion.  
PT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270DBC4E65 CRC64;  
Q957S8 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSIRH GREKKRRTDR TRA

! ! AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 20 AA.

ID Q9T2K6  
AC Q9T2K6;  
DT 01-MAY-2000 (TREMBrel. 13, Created)  
DT 01-JUN-2000 (TREMBrel. 14, Last sequence update)  
RT Abies alba.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DDAJ databases.  
DR EMBL: AE359456; AAK48935.1; -.

KW Mitochondrion.  
PT NON\_TER 22 22  
SQ SEQUENCE 22 AA; 2711 MW; 97BF270DBC4E655 CRC64;  
Q957T4 Length: 22 December 22, 2002 19:30 Type: P Check: 9875 ..

1 MPTSNQSIRH GREKKRRTDR TRA

! ! AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 23 AA.

ID Q957T0  
AC Q957T0;  
DT 01-DEC-2001 (TREMBrel. 19, Created)  
DT 01-JUN-2002 (TREMBrel. 21, Last annotation update)

DE Ribosomal protein S12 (Fragment).  
GN RPS12.  
OS Pinus mugo.  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus;  
OX NCBI\_TaxID=20528;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wachowiak W., Liepelt S., Prus-Glowacki W.;  
RT "Sequence analysis of chosen regions of the mitochondrial genome of  
RT *Pinus mugo* and *P. sylvestris*.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DDAJ databases.  
DR EMBL: AE359576; AAK39117.1; -.  
KW Mitochondrion.  
PT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270DBC4E65 CRC64;

Q9T2K6 Length: 20 December 22, 2002 19:30 Type: P Check: 5922 ..

1 GSSSLADRXLGSLAXDGES

! ! AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 24 AA.

ID Q9T2H3  
AC Q9T2H3;  
DT 01-MAY-2000 (TREMBrel. 13, Created)  
DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBrel. 16, Last annotation update)

DE Chaperonin 21 (Fragment).  
OS Spinatina oleracea (Spinacia).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;

RN [1]

RP SEQUENCE  
 RX MEDLINE=95394976; PubMed=7665625;  
 RA Ryan M.T.; Maylor D.J.; Hoogendoord N.J.; Hoj P.B.;  
 RRL J. Biol. Chem. 270:22037-22043(1995);  
 SEQID 24 AA; 2642 MW; B79781E1005A51A CRC64;  
 SQ 092H3 Length: 24 December 22, 2002 19:30 Type: P Check: 2853 ..

1 ATVVAKPTS IKPTADRLI KIKE

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 18 AA.

ID Q37852; AC 037852; DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE A protein (Fragment).

OS Bacteriophage R17.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;  
 OC Levivirus.  
 OX NCBI\_TaxID=12026;

RN [1]

RP SEQUENCE FROM N\_A.  
 RX MEDLINE=74126852; PubMed=361645;  
 RA Renning U.F.E.; Coulson A.; Schoenmakers J.G.G.;  
 RT "A sequence of 54 nucleotides from the A-protein cistron of coliphage-  
 R17 RNA"; Eur. J. Biochem. 41:431-438(1974).  
 DR EMBL; M24814; AA32171; ..

FT NON\_TER 18 AA; 18 ..

FT NON\_TER 18 AA; 2217 MW; CB3DE9FD8852F11E CRC64;

SQ 037852 Length: 18 December 22, 2002 19:30 Type: P Check: 3123 ..

1 ALRVLAALNED RKFRSKHV

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 28 AA.

ID Q42209; AC 042209; DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Ubiquitin conjugating enzyme UBC10 (Fragment).

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicots; Rosidae;  
 OC eurosid I; Cucurbitales; Cucurbitaceae; Cucumis;  
 OX NCBI\_TaxID=3659;

RN [1]

RP SEQUENCE FROM N\_A.  
 RX TISSUE-COTYLEDON;  
 MEDLINE=96104305; PubMed=8564304;  
 RA Toyama T.; Teramoto H.; Takeya G.; Tsuji H.;  
 RT "Cytokinin induces a rapid decrease in the levels of mRNAs for  
 catalase, 3-hydroxy-3-hydroxyglutaryl CoA reductase, lectin and other  
 unidentified proteins in etiolated cotyledons of cucumber";  
 RL Plant Cell Physiol. 36:1349-1359(1995).  
 DR EMBL; D63385; BA09701; ..

FT NON\_TER 1 1 ..

SQ 039633 Length: 23 December 22, 2002 19:30 Type: P Check: 1444 ..

1 NGSQADRSVG QKLAPHLNVR PSI

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 26 AA.

ID P82195; AC P82195; DT 01-JUN-2000 (TREMBLrel. 14, Created)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L1B beta (Fragment).  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicots;  
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia;  
 OX NCBI\_TaxID=3562;

RN [1]

RP SEQUENCE  
 RX STRAIN=CV; ALWARO; TISSUE=LEAF;  
 MEDLINE=20435793; PubMed=10974046;  
 RA Yamaguchi K.; Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -1 - FUNCTION: THIS PROTEIN BANDS DIRECTLY TO 5S RIBOSOMAL RNA.  
 CC -1 - SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -1 - TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -1 - MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 KDA.  
 CC -1 - SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.

! !AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 29 AA.

ID Q9FP4; AC Q9FP4; DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Genomic DNA, chromosome 5, PI clone: MVP2.

KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON\_TER 26 AA; 3184 MW; C5664EC1B2AB7C97 CRC64;  
 SQ SEQUENCE 26 AA: ..

P82195 Length: 26 December 22, 2002 19:30 Type: P Check: 7110 ..

1 KATTRREDRT ARHVIRKKV EGTPER  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 29 AA.

ID P82196 PRELIMINARY; PRT; 29 AA.

DT 01-JUN-2000 (TREMBrel. 14, Created)  
 DT 01-JUN-2000 (TREMBrel. 14, Last sequence update)  
 DT 01-MAR-2001 (TREMBrel. 16, Last annotation update)

DE Chloroplast 50S ribosomal protein L18 alpha (Fragment).

OS Spinacia olaracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.

NCBI\_TaxID=3562;

[1]

SEQUENCE:  
 RC STRAIN=CV\_ALWARO; TISSUE=LEAF;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).

CC -!- FUNCTION: THIS PROTEIN Binds DIRECTLY TO 5S RIBOSOMAL RNA.

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.

CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 kDa.

CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.

KW Ribosomal protein; chloroplast; rRNA-binding.

FT NON\_TER 29 AA; 3117 MW; 86207D5564EC1B2 CRC64;

SQ SEQUENCE 29 AA: ..

P82196 Length: 29 December 22, 2002 19:30 Type: P Check: 4154 ..

1 KAHPREDRT ARHVRKVV EGTPERXXL

!!AA\_SEQUENCE 1.0  
 ID P83089 PRELIMINARY; PRT; 19 AA.

DT 01-OCT-2001 (TREMBrel. 18, Created)  
 DT 01-OCT-2001 (TREMBrel. 18, Last sequence update)

DE Thylakoid lumenal 25.3 kDa protein (Fragment).

OS Spinacia olaracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.

NCBI\_TaxID=3562;

[1]

RN SEQUENCE, SUBCELLULAR LOCATION, AND MASS SPECTROMETRY.

RC TISSUE=LEAF;  
 RX PubMed=11719511;

Schubert M., Petersson U.A., Haas B.J., Funk C., Schroeder W.P.,  
 Kieselbach T.;  
 "Proteome map of the chloroplast lumen of Arabidopsis thaliana.";  
 J. Biol. Chem. 277:8354-8355(2002).

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN.

CC -!- MASS SPECTROMETRY: MW=25300; METHOD=MALDI.

KW Chloroplast; Thylakoid.

FT NON\_TER 19 AA; 2093 MW; EC33422F8633F1F9 CRC64;

SQ SEQUENCE 19 AA: ..

P83089 Length: 19 December 22, 2002 19:30 Type: P Check: 4307 ..

1 AIAA\_PLOT TTDYRFFD

!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 29 AA.

ID 0958D2 PRELIMINARY; PRT; 16 AA.

AC Q9S8D2; 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)

DE CMTI-B-TRYPSIN inhibitor.

OS Cucumis melo (Muskmelon).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosid I; Cucurbitales; Cucurbitaceae; Cucumis.

NCBI\_TaxID=3656;

RN [1]

RP SEQUENCE.  
 RX MEDLINE=96015144; PubMed=8537309;

RA Lee C.F., Lin J.Y.;

RT "Amino acid sequences of trypsin inhibitors from the melon Cucumis  
 melo.";  
 J. Biochem. 118:18-22(1995).

RL HSSP; P12071; 2E11.

DR Interpro; IPR000737; Squash.

DR Pfam; PF00299; squash\_1.

DR PRINTS; PR0293; SQUASHINHTR.

DR PRODOM; PD03401; Squash; 1.

DR SMART; SM00286; PRT; 1.

DR PROSITE; PS00286; SQUASH\_INHIBITOR; 1.

SQ SEQUENCE 29 AA: 3196 MW; A4BCFF7A1AC300B CRC64;

Q9S8D2 Length: 29 December 22, 2002 19:30 Type: P Check: 2605 ..

1 VGCPRILMKC KTDRCCLTG CTKRNGYCG

!!AA\_SEQUENCE 1.0  
 ID Q9S8D8 PRELIMINARY; PRT; 16 AA.

AC Q9S8D8; 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMBrel. 14, Last annotation update)

DT 01-JUN-2000 (TREMBrel. 14, Last annotation update)

DE 23 kDa heat-induced protein (Fragment).

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; eustreerids I; Solanales; Solanaceae; Solanum.

NCBI\_TaxID=4081;

RN [1]

RP SEQUENCE.  
 RX MEDLINE=96351184; PubMed=8742333;

RA Sabehat A., Weiss D., Lurie S.;

RT "The correlation between heat-shock protein accumulation and  
 persistence and chilling tolerance in tomato fruit.";  
 Plant Physiol. 110:531-537(1996).

SQ SEQUENCE 16 AA: 1840 MW; 9B0D3F7BFEE6CE3 CRC64;

Q9S8D8 Length: 16 December 22, 2002 19:30 Type: P Check: 484 ..

1 NTNTQMTAYD QDDRGX

!!AA\_SEQUENCE 1.0  
 ID Q9QF2 PRELIMINARY; PRT; 12 AA.

AC Q9QF2; 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)

DE 01-JUN-2002 (TREMBrel. 21, Last annotation update)

DT TRANSFERRIN-PEPTIDE 21 (Fragment).

DE Rattus sp.;  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI\_TaxID=10118;

RN [1]

RP SEQUENCE.  
 RX MEDLINE=2165927; PubMed=179188;

RA Cavanaugh P.G., Nicolson G.L.;

RT "Lung-derived growth factor that stimulates the growth of lung-  
 metastasizing tumor cells: identification as transferrin.";  
 J. Cell. Biochem. 47:261-271(1991).

FT NON\_TER 1 1  
 FT NON\_TER 12 12  
 SEQUENCE 12 AA; 1436 MW; 40AD1DFA420AAD3 CRC64;  
 Q9QVF2 Length: 12 December 22, 2002 19:30 Type: P Check: 5863 ..

1 NLANKADRDQ YE

!AA\_SEQUENCE 1.0  
 PRELIMINARY; PRT; 18 AA.

ID Q9CINO  
 AC Q9CINO;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE B230218P12Rik protein (Fragment).  
 GN B230218P12Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N\_A.

RC STRAIN=C57BL6J; TISSUE=CORPORA QUADRIGEMINA;  
 RX MEDLINE=2108660; PubMed=1121851;

RA Kawai J., Shinagawa K., Shiba T., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Arizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuoka Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustinich C., Hill D., Hoffmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Maslima J., Mazzarelli J.J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasai H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection." Nature 409:685-690(2000).

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shihata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection." Nature 409:685-690(2000).

DR EMBL; AK021007; BAB32275\_1;  
 DR MGD; MGI:1926130; B230218P12Rik.  
 FT NON\_TER 18 18  
 SEQUENCE 18 AA; 2136 MW; 72A0BET72BCA48C8 CRC64;

Q9CINO Length: 18 December 22, 2002 19:30 Type: P Check: 2989 ..

1 MPVKKKDTPR ALSLLEPY

!AA\_SEQUENCE 1.0  
 PRELIMINARY; PRT; 27 AA.

ID Q9C332;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DE 261002005Rik protein (Fragment).  
 GN B2302005Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N\_A.

RC STRAIN=C57BL6J; TISSUE=EMBRYO; MEDLINE=2108660; PubMed=1121851;

RA Kawai J., Shinagawa A., Shiba T., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hoffmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Maslima J., Mazzarelli J.J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shihata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection." Nature 409:685-690(2000).

DR EMBL; AK01213; BAB31604\_1; ..  
 MGD; MGI:1914299; 2610020J05Rik.

NON\_TER 27 27  
 SEQUENCE 27 AA; 3299 MW; 95C1A07254A2C489 CRC64;

Q9CS32 Length: 27 December 22, 2002 19:30 Type: P Check: 8838 ..

1 MAENTDRNQI EKLINRVLK EQLVER

!AA\_SEQUENCE 1.0  
 PRELIMINARY; PRT; 25 AA.

ID Q9CS32;  
 AC Q9CS32;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

RA Rettus norvegicus (Rat);  
 OS Rettus norvegicus (Rat);  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus. OX NCBI\_TaxID=10116;  
 RN [1]

RP SEQUENCE FROM N\_A.

RC STRAIN=WISTAR;

RX TISSUE=BONE MORPHOGENETIC PROTEIN-IMPLANTED SUBCUTANEOUS MUSCLE; MEDLINE=94271262; PubMed=7911662;

RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;  
 RT "Changes in homeobox-containing gene expression during ectopic bone formation induced by bone morphogenic protein.";  
 RL Biochem Biophys Res Commun. 201:980-987(1994).  
 DR EMBL; S7126; AB31006\_1; ..  
 DR Titterpro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1;  
 DR PRODOM; PD00010; Homeobox; 1;  
 DR DNA-binding; Homeobox; Nuclear protein.  
 FT NON\_TER 1 1  
 SEQUENCE 25 AA; 3111 MW; 008AA81607BD2203 CRC64;

Q9CS32 Length: 25 December 22, 2002 19:30 Type: P Check: 5506 ..

1 LFNMVLTTR RLEISKTNL TDQV

!AA\_SEQUENCE 1.0  
 PRELIMINARY; PRT; 33 AA.

ID Q98226;  
 AC Q98226;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DR MS2F76 (Fragment).  
 GN MS2F76.

OS Mus musculus (Mouse);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. OX NCBI\_TaxID=10090;  
 RN [1]  
 SEQUENCE FROM N\_A.

RC STRAIN=C57BL6;

RX MEDLINE=98296253; PubMed=9530514;  
 RA Agata Y.; Matzuda E.; Shimizu A.;  
 RT "Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc  
 finger proteins by degenerate PCR. ";  
 RL Gene 21:55-60(1998);  
 DR EMBL; AB010321; BA31377.1; -.  
 DR Interpro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF0096; zf-C2H2\_1.  
 DR PRODOM; P00003; Znf\_C2H2\_1.  
 DR SMART; SM0355; Znf\_C2H2\_1.  
 DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 1.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 33 33  
 SQ SEQUENCE 33 AA; 3730 MW; D8D31E5B5C54C538 CRC64;  
 088226 Length: 33 December 22, 2002 19:30 Type: P Check: 2429 ..  
 1 IHTGKPKYRC AEGKAKFTDR SNLIKHQTH TGE  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 32 AA.  
 ID Q9J100; AC Q9J100; DT 01-OCT-2000 (TREMBrel. 15, Created)  
 DT 01-JUN-2002 (TREMBrel. 15, Last sequence update)  
 DE Protocadherin-T5 (Fragment).  
 GN PCDH-T5  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.; STRAIN=FISCHER44; TISSUE=TESTIS;  
 RX MEDLINE=20114371; PubMed=1065049;  
 RA Johnson K.J.; Patel S.R.; Boekelheide K.;  
 RT "Multiple cadherin superfamily members with unique expression profiles  
 are produced in rat testis";  
 RT Endocrinology 141:675-683(2000).  
 DR EMBL; AF176769; AFM87074.1; -.  
 DR Interpro; IPR002126; cadherin\_1.  
 DR Pfam; PF00028; cadherin\_1.  
 FT NON\_TER 1 1  
 FT NON\_TER 32 32  
 SQ SEQUENCE 32 AA; 3369 MW; A6159FED42844051 CRC64;  
 Q9J100 Length: 32 December 22, 2002 19:30 Type: P Check: 1050 ..  
 1 ALQAFEFHVQ ATDRGSPALS SQALRVVVL DN  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 25 AA.  
 ID Q9QX46; AC Q9QX46; DT 01-MAY-2000 (TREMBrel. 13, Created)  
 DT 01-MAY-2000 (TREMBrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE Amyloid protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.; STRAIN=129/SVJ; TISSUE=LIVER;  
 RX MEDLINE=20102720; PubMed=10336908;  
 RA Bergsorf C.; Palica K.; Krieger S.; Masters C.L.; Beyreuther K.;  
 RT "Identification of cis Elements Regulating Exon 15 Splicing of the  
 Amyloid Precursor Protein Pre-mRNA. ";  
 RT J. Biol. Chem. 275:2046-2056(2000).  
 DR EMBL; AF199005; AFM20194.1; -.  
 DR EMBL; AF199003; AFM20194.1; JOINED.

DR EMBL; AF199004; AAF20194.1; JOINED.  
 FT NON\_TER 1 1  
 DT 01-NON-1996 (TREMBrel. 01, Created)  
 DT 01-MAR-2001 (TREMBrel. 16, Last annotation update)  
 SQ SEQUENCE 25 AA; 2608 MW; 276F026D582421D0 CRC64;  
 09QX46 Length: 25 December 22, 2002 19:30 Type: P Check: 4770 ..  
 1 ENEVEPVNDAR PAADQLITR PGSSL  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 20 AA.  
 ID Q9E700; AC Q9E700; DT 01-MAR-2001 (TREMBrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBrel. 16, Last annotation update)  
 DE EIF4H (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.; STRAIN=29/SV;  
 RA Green E.D.; DR Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF289665; AAC99335.1; -.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2188 MW; 8AFFACOFFB953EBF5 CRC64;  
 Q9E700 Length: 20 December 22, 2002 19:30 Type: P Check: 5807 ..  
 1 MADFDTYDDR AVSSFFGGRG  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 19 AA.  
 ID Q9Z115; AC Q9Z115; DT 01-MAY-1999 (TREMBrel. 10, Created)  
 DT 01-MAY-1999 (TREMBrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBrel. 10, Last annotation update)  
 DE Calcineurin A alpha (EC 3.1.3.16) (Fragment).  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE FROM N.A.; STRAIN=LIVER;  
 RX TISSUE=LIVER;  
 RA MEDLINE=93111954; PubMed=1335233;  
 RA Chang C.; Takechi T.; Mukai H.; Shuncho H.; Kuno T.; Tanaka C.;  
 RT "Molecular cloning and characterization of the promoter region of the  
 calcineurin A alpha gene";  
 RT Biochem. J. 288:801-805(1992).  
 DR EMBL; D10480; BA01283.1; -.  
 KW Hydrolase.  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 2115 MW; F905FB134CCFC57 CRC64;  
 Q9Z115 Length: 19 December 22, 2002 19:30 Type: P Check: 4927 ..  
 1 MSEKKAIPK LSTTDWKV  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 18 AA.  
 ID Q62256; AC Q62256; DT 01-NOV-1996 (TREMBrel. 01, Created)  
 DT 01-NOV-1996 (TREMBrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
 DE Spermatogenic-specific proenkphalin.  
 GN PENK-RS.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;

[1] RP SEQUENCE FROM N.A.  
 RX MEDLINE=20281163; PubMed=2355920;  
 RA Kilpatrick D.L., Zinn S.A., Fitzgerald M., Higuchi H., Sabol S.L.,  
 RA Meyerhardt J.;  
 RT "transcription of the rat and mouse proenkcephalin genes is initiated  
 at distinct sites in spermatogenic and somatic cells.";  
 RT Mol. Cell. Biol. 10:3717-3726(1990).  
 DR MGI: M5181; AA:A40127-1; -;  
 DR MGD: MGI:104628; PenkTS;  
 SQ SEQUENCE 18 AA; 2043 MW; B96E10CC7049FA76 CRC64;  
 062256 Length: 18 December 22, 2002 19:30 Type: P Check: 3193 ..

1 MSSGKQDSPW EDTRPPGR

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 25 AA.  
 ID Q63985; AC 063985; DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Aristaleless related homeobox protein Arx (Fragment).  
 OS Cricetulus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetus. [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=4271262; Pubmed=911662;  
 RA Tamura T., Olda S., Takada K., Maruoka Y., Sasaki S.;  
 RT "Changes in homeobox-containing gene expression during ectopic bone  
 formation induced by bone morphogenetic protein";  
 RL Biochem. Biophys. Res. Commun. 201:980-987(1994).  
 EMBL: S1275; AA:31001..1; -;  
 DR InterPro: IPR001556; Homeobox.  
 DR Pfam: PF00046; Homeobox; 1.  
 DR ProDom: PD000010; Homeobox; 1.  
 KW DNA binding; Homeobox; Nuclear protein.  
 FT NON\_TER 1  
 SO SEQUENCE 25 AA; 3134 MW; C54EB951607BD27B1 CRC64;  
 063985 Length: 25 December 22, 2002 19:30 Type: P Check: 5721 ..

1 LFNNYLTRER RUEISRSVHL TDROV

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 32 AA.  
 ID Q61159; AC 061159; DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE NNB-AG (Fragment).  
 RRG1 OR NNB-AG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NNB-AG (fragment). [1]  
 RR SEQUENCE FROM N.A.  
 RX MEDLINE=96315660; PubMed=8754834;  
 RA Shago M., Glaguere V.;  
 RT "Isolation of a novel retinoic acid-responsive gene by selection of  
 genomic fragments derived from CPG island-enriched DNA.";  
 RT Mol. Cell. Biol. 16:4337-4348(1996).  
 DR MGD: MGI:108048; Rrg1;  
 DR NON\_TER 32; 32;  
 SQ SEQUENCE 32 AA; 3470 MW; 9571FBDB80B1FCAB4 CRC64;

061159 Length: 32 December 22, 2002 19:30 Type: P Check: 837 ..

1 MAASMDVFS FCYGVADRAR GSVEVRYVDs IK

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 40 AA.  
 ID Q923H1; AC Q923H1; DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Aristaless related homeobox protein Arx (Fragment).  
 OS Cricetulus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetus. [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=92207982; Pubmed=1372829;  
 RA Nakashima N., Miyazaki K., Ishikawa M., Yatohige T., Ogawa H.,  
 RA Uchibori H., Matsumoto I., Seno N., Hayashi M.;  
 RT "Vitronectin domain in evolution but uniformity in ligand binding  
 and size of the core polypeptide";  
 RT Bloch. Biophys. Acta 1120:1-10(1992).  
 RL NON\_TER 1

FT	NON_TER	24	24	MW;	0CD91505B6F22D0A CRC64;
SQ	SEQUENCE	24 AA:	2736	MW;	0CD91505B6F22D0A CRC64;
1	XXESXKGRT EGFnADRKKQ XXEL				
! 1AA_SEQUENCE 1.0 PRELIMINARY; PRT; 32 AA.					
ID_09QV89; Q9QV89; 01-MAY-2000 (TREMBLrel. 13, Created)	PRT;	32 AA.			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)					
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)					
DR Sip-14-Fatty acid-binding protein (Fragment).					
OS Rattus sp.					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. NCBI_TaxID=10118; [1]					
RN RX	SEQUENCE				
RP B0319171; PubMed=8447836;					
RA Bansal M.P.; Medina D.;					
RT "Expression of fatty acid-binding proteins in the developing mouse mammary gland.", Biochem. Biophys. Res. Commun. 191:61-69(1993).					
RL HSSP; P05413; IMLT.					
DR InterPro: IPR00566; Lipocin_cytFABP.					
PFam: PF00061; Lipocalin_1.					
DR NON_TER 1					
FT NON_TER 1					
FT NON_TER 32		32			
FT SEQUENCE 32 AA;		3495 MW;	67D23A0146E46AFE CRC64;		
Q9QV89 Length: 32 December 22, 2002 19:30 Type: P Check: 581 ..					
1 EISFQLVGEF DEVTAADDRKV KSYVTLGGK LV					
! 1AA_SEQUENCE 1.0 PRELIMINARY; PRT; 17 AA.					
ID_09QV71; Q9QV71; 01-MAY-2000 (TREMBLrel. 13, Created)	PRT;	17 AA.			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)					
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)					
DR Lactase-phlorizin hydrolase (Fragment).					
OS Rattus sp.					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. NCBI_TaxID=10118; [1]					
RN RX	SEQUENCE				
RP MEDLINE=9329388; PubMed=8514793;					
RA Dudley M.A.; Hatchey D.L.; Quarchi A.; Hutchens T.W.; Nichols B.I.,					
RA Rosemberger J.; Perkinsom J.S.; Cook G.; Reeds P.J.;					
RT "In vivo sucrase-isomaltase and lactase-phlorizin hydrolase turnover in the fed adult rat.", J. Biol. Chem. 288:13609-13616(1993); SEQUENCE 17 AA; 1947 MW; 08941FD316F9592F CRC64; [1]					
Q9QV71 Length: 17 December 22, 2002 19:30 Type: P Check: 1363 ..					
1 DWEDBRNFIAA GPLTNLDL					
! 1AA_SEQUENCE 1.0 PRELIMINARY; PRT; 25 AA.					
ID_061461; Q61461; 01-NOV-1996 (TREMBLrel. 01, Created)	PRT;	25 AA.			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DT 01-MAR-2002 (TREMBLrel. 09, Last annotation update)					
DR Cytochrome P-450b (Phenobarbital-inducible) (Fragment). CYP2B13.					
OS Mus musculus (Mouse).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI_TaxID=10090; [1]					
RN RX	SEQUENCE FROM N.A.				
RP MEDLINE=84207435; PubMed=5547088;					
RA Stupans I., Ikeda T., Kessner D.J., Nebert D.W.;					"Characterization of a cDNA clone for mouse phenobarbital-inducible cytochrome P-450b."
RT DNA 3:129-137(1984).					
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.					
DR InterPro: IPR001126; Cytochrome_P450.					
DR Pfam: PF00067; P450; 1.					
DR EMBL: K02409; AA351501; 1.					
DR MED; MGI:88599; Cyp2b13.					
FT NON_TER 1					
FT NON_TER 25		25			
SQ SEQUENCE 25 AA; 2979 MW; B9DFF7A57355A1B0 CRC64; 061461 Length: 25 December 22, 2002 19:30 Type: P Check: 4925 ..					
1 SHRPLTLLDR SKMPYTDAVI HEIQR					
! 1AA_SEQUENCE 1.0 PRELIMINARY; PRT; 19 AA.					
ID_05747; Q5747; 01-NOV-1996 (TREMBLrel. 01, Created)	PRT;	19 AA.			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DT 01-NOV-2001 (TREMBLrel. 17, Last annotation update)					
DR Outer coat protein (VP2) (Fragment).					
OS Blue-tongue virus.					
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus. NCBI_TaxID=12591; [1]					
RN RX	SEQUENCE FROM N.A.				
RP MEDLINE=88265863; PubMed=2838961;					
RA Gould A.R., Hyatt A.D., Eaton B.T.;					
RT "Morphogenesis of a blue-tongue virus variant with an amino acid alteration at a neutralization site in the outer coat protein, VP2."					
RL Virology 165:23-32(1988); EMU; M2135; AA04845; [1]					
DR InterPro: IPR001742; Orbi_VP2.					
DR Pfam: PF00898; Orbi_VP2; 1.					
KW Coat Protein.					
FT NON_TER 1		1			
FT NON_TER 19		19			
SQ SEQUENCE 19 AA; 2608 MW; F0428841265D3061 CRC64; 065747 Length: 19 December 22, 2002 19:30 Type: P Check: 4484 ..					
1 ERKIFHHRN QRQBEDDRFY					
! 1AA_SEQUENCE 1.0 PRELIMINARY; PRT; 35 AA.					
ID_065538; Q65538; 01-NOV-1996 (TREMBLrel. 01, Created)	PRT;	35 AA.			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)					
DR 3' proximal protein (Fragment).					
OS Ebola virus.					
OC Virtues; ssRNA negative-strand viruses; Mononegavirales; Filoviridae. Ebola-like viruses. NCBI_TaxID=11268; [1]					
RN RX	SEQUENCE FROM N.A.				
RP MEDLINE=86124724; PubMed=3946083;					
RA Kiley M.P., Wilusz J., McCormick J.B., Keene J.D.;					
RT "Conservation of the 3' terminal nucleotide sequences of Ebola and Marburg virus.", Virology 149:251-254(1986); EMU; M3062; AA042976; 1.					
RT DR					
FT NON_TER 35		35			
SQ SEQUENCE 35 AA; 4142 MW; F7778E5FC6283092 CRC64; Q66538 Length: 35 December 22, 2002 19:30 Type: P Check: 8137 ..					

1 MRKINNFLSL KFDDRLNLKLK LLICNHTVDS EPHNS  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 9 AA.  
 ID Q83622; AC Q83622; DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical 1.1 kDa protein (Fragment).  
 OS Murray valley encephalitis virus (Fragment).  
 OC viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Flavivirus.  
 NCBI\_TAXID=11079;  
 [1]  
 SEQUENCE FROM N.A.  
 RP MEDLINE=88118912; PubMed=2826633;  
 RA Hahn C.S., Hahn Y.S., Rice C.M., Dee E., Dalgarno L., Strauss E.G.,  
 RA Strauss J.H.;  
 RT "Conserved elements in the 3' untranslated region of flavivirus RNAs  
 and potential cyclization sequences.";  
 RT J. Mol. Biol. 188:33-41(1987).  
 RL EMBL: M35172; AAA66627.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SEQUENCE 9 AA; 1055 MW; FF36D40AB05A2C1 CRC64;  
 SQ 083622 Length: 9 December 22, 2002 19:30 Type: P Check: 3517 ..

1 THVSEDRVL  
 RP !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 22 AA.  
 ID Q84254; AC Q84254; DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1998 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE LBXL protein (Fragment).  
 GN Gallus gallus (Chicken).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae  
 OC Gallius.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dietrich S., Schubert F.R., Healy C., Sharpe P.T., Lumsden A.;  
 RT "Specification of the hypaxial musculature";  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Schubert F.R., Dietrich S., Chapman S.C., Lumsden A.;  
 RT "Expression of the Lbx1 gene in the chick embryo";  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: AJ22831; GCA1016.1; -.  
 DR InterPro: IPR01156; Homeobox.  
 DR Pfam: PF0045; homeobox; 1.  
 DR PRODom: PD00010; Homeobox; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 29 AA; 3352 MW; CABBC02415F1D4F7 CRC64;

084254 Length: 22 December 22, 2002 19:30 Type: P Check: 9066 ..

1 NAGPKPGTP EDVADRRPDL PE  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 39 AA.  
 ID Q90237; AC Q90237; DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Homeobox (Fragment).  
 OS Carassius auratus (Goldfish); Craniata; Vertebrata; Euteleostomi;  
 OC Cyprinidae: Carassius.  
 OC Cyprinidae: Carassius.  
 RN NCBI\_TAXID=7057;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=RETINA;  
 MEDLINE=93121357; PubMed=8096640;  
 RA Levine E.M., Schechter N.;  
 RT Homeobox genes are expressed in the retina and brain of adult  
 goldfish.;"  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).  
 DR EMBL: L09693; AAC9178.1; -.  
 DR InterPro: IPR01336; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.

093438 Length: 29 December 22, 2002 19:30 Type: P Check: 3493 ..

1 RFLIQQYKLS PARDQIAQL GLTRAQVTT  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
 ID Q90237; AC Q90237; DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Homeobox (Fragment).  
 OS Carassius auratus (Goldfish); Craniata; Vertebrata; Euteleostomi;  
 OC Cyprinidae: Carassius.  
 OC Cyprinidae: Carassius.  
 RN NCBI\_TAXID=7057;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=RETINA;  
 MEDLINE=93121357; PubMed=8096640;  
 RA Levine E.M., Schechter N.;  
 RT Homeobox genes are expressed in the retina and brain of adult  
 goldfish.;"  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).  
 DR EMBL: L09693; AAC9178.1; -.  
 DR InterPro: IPR01336; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.

DR PRODOM; P000010; Homeobox; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 33338 MW; 55318651A81607BD CRC64;

Q92297 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1 LFNMYLTER RLEISKSINL TDROVKI

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.

ID Q92298  
 AC Q92298;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Homeobox (Fragment).

OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Carassius.

RN [1]

SEQUENCE FROM N.A.

RC TISSUE-RETINA;  
 RX MEDLINE=93219357; PubMed=8096640;  
 RA Levine E.M.; Schechter N.;  
 RT "Homeobox genes are expressed in the retina and brain of adult goldfish.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).  
 DR EMBL: L09694; AAA49179.1;  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS: PR00024; HOMEBOX.  
 DR PRODOM; PD000010; Homeobox; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 3375 MW; 52B7254F951607BD CRC64;

Q92298 Length: 27 December 22, 2002 19:30 Type: P Check: 9642 ..

1 LFNMYLTER RLEISRSVHL TDROVKI

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.

ID Q92295  
 AC Q92295;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE Homeobox (Fragment).

OS Petromyzon marinus (Sea lamprey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia; Petromyzontiformes; Petromyzontidae; Petromyzon.

RN NCBI\_TaxID=7777;

RP SEQUENCE FROM N.A.

RX MEDLINE=93317699; PubMed=8101001;

RA Pendleton J.W.; Nagai B.K.; Murtha M.T.; Ruddle F.H.;  
 RT "Expansion of the Hox gene family and the evolution of chordates.";  
 RL Pandelton J.W.; Nagai B.K.; Murtha M.T.; Ruddle F.H.;  
 RT "Expansion of the Hox gene family and the evolution of chordates.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).

DR EMBL: L14900; AAA02535.1; -.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS: PR00024; HOMEBOX.

FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 3319 MW; 53CD02F1671607BD CRC64;

Q07150 Length: 27 December 22, 2002 19:30 Type: P Check: 9407 ..

1 LFSMLITRER RLEISHLLSI TDROVKI

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 14 AA.

ID Q07150  
 AC Q07150;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE Homeobox (Fragment).

OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianinae; Gallus.  
 OC NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN-WHITE LEGHORN;

RA MEDLINE=98111813; PubMed=9473273;

RT Peale F.V.; Mason K.; Hunter A.W.; Bothwell M.;  
 RT "Multiplex display polymerase chain reaction amplifies and resolves related sequences sharing a single moderately conserved domain.";  
 RL Anal. Biochem. 256:158-160(1998).

DR EMBL: U34614; AAC36453.1; -.

FT NON\_TER 1 1  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1673 MW; 81258FC9E81FDA45 CRC64;

Q073591 Length: 14 December 22, 2002 19:30 Type: P Check: 8165 ..

1 RLEISKSINL TDROVKI

!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 13 AA.

ID Q073591  
 AC Q073591;  
 DT 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Phosholenman (FFYD domain-containing ion transport regulator 1) (Fragment).

OS Lamna nasus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Galeomorphi; Galeoidea; Lamniformes; Alopiidae; Lamna.  
 OC NCBI\_TaxID=7849;

RN [1]

1 LFNMLITRER RLEISRGVNL TDROVKI

RP SEQUENCE.  
 RC ISSUE=RECTAL GLAND;  
 RA Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.E.;  
 RT "Phospholemman in the rectal gland of sharks.";  
 RL Submitted (JUN-2001) to the SWISS-PROT data bank;  
 CC -!- FUNCTION: A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT  
   WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN  
   MUSCLE CONTRACTION.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSOMAL  
   MEMBRANE.  
 CC -!- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN  
   KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT  
   TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC  
   STIMULATION (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE FXYD FAMILY.  
 DR InterPro; IPR000272; ATPG1\_LM\_MAT8.  
 DR PROSITE; PS01310; FXID; PARTIAL.  
 KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;  
 KW Microsome.  
 FT NON\_TER 13 AA; 1542 MW; 081373C69724A444 CRC64;  
 SQ SEQUENCE 13 AA; ..

P83009 Length: 13 December 22, 2002 19:30 Type: P Check: 7125 ..

1 VSDVPNNDDR FTY

!!AA\_SEQUENCE 1.0  
 ID P83010 PRELIMINARY; PRT; 13 AA.  
 AC P83010;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Phospholemman (FXYD domain-containing ion transport regulator 1)  
 DE (Fragment).  
 OS Triakis scyllium (Leopard shark) (Triakis scyllia).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;  
 OC Triakis.  
 OX NCBI\_TaxID=30494;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=RECTAL GLAND;  
 RA Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.E.;  
 RT "Phospholemman in the rectal gland of sharks.";  
 RL Submitted (JUN-2001) to the SWISS-PROT data bank;  
 CC -!- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT  
   WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN  
   MUSCLE CONTRACTION.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSOMAL  
   MEMBRANE.  
 CC -!- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN  
   KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT  
   TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC  
   STIMULATION (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE FXID FAMILY.  
 DR InterPro; IPR000272; ATPG1\_LM\_MAT8.  
 DR PROSITE; PS01310; FXID; PARTIAL.  
 KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;  
 KW Microsome.  
 FT NON\_TER 13 AA; 1497 MW; 1D14C92C1AE144 CRC64;  
 SQ SEQUENCE 13 AA; ..

P83010 Length: 13 December 22, 2002 19:30 Type: P Check: 6920 ..

1 AGEPANNEDR FNY

!!AA\_SEQUENCE 1.0  
 ID Q90263 PRELIMINARY; PRT; 27 AA.  
 AC 090263;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hoxcilla (Fragment).  
 RN [1]

RN SEQUENCE FROM N.A.

RP RT "Isolation of BAC Clones That Cover Seven Hox Clusters in Medaka Genome.";  
 RN RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB055740; BAB62853; 1; ..  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR ProDom; PD000110; Homeobox; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 3368 MW; 548C73EA983E9158 CRC64;

Q902G3 Length: 27 December 22, 2002 19:30 Type: P Check: 9516 ..

1 FFNVYINKER RIQLSRMLN TDROVRI

!!AA\_SEQUENCE 1.0  
 ID Q9PS38 PRELIMINARY; PRT; 20 AA.  
 AC Q9PS38;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE PO GLYCOPROTEIN-MYELIN major structural protein (Fragment).  
 OS Rana catesbeiana (Bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Randiae; Rana.  
 OX NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93002191; PubMed=1382532;  
 RA Karthigasan J., Bauer T.K., Teplow D.B., Saavedra R.A.,  
 RA Kirschner D.A.;  
 RT "Phylogenetically conserved amino acids of MBP and P0 from amphibian myelin";  
 RL J. Mol. Neurosci. 3:185-188(1992).  
 FT NON\_TER 1 1  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2351 MW; AB238ACED7837676 CRC64;

Q9PS38 Length: 20 December 22, 2002 19:30 Type: P Check: 6634 ..

1 IEVYTDRBIO SVNGSKVHLV

!!AA\_SEQUENCE 1.0  
 ID Q9PS32 PRELIMINARY; PRT; 23 AA.  
 AC Q9PS32;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Nonmuscle alpha actinin 115 kDa isoform (Fragment).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=9300311; PubMed=1334489;  
 RA Imamura M.; Masaki T.;  
 RT "A novel nonmuscle alpha actinin. Purification and characterization of  
   chicken lung alpha-actinin.";  
 RL J. Biol. Chem. 267:25927-25933(1992).  
 FT NON\_TER 1 1  
 FT NON\_TER 23 23

SQ SEQUENCE 23 AA; 2823 MW; EBED28437F668363 CRC64;  
 Q9PS32 Length: 23 December 22, 2002 19:30 Type: P Check: 1521 ..

1 LASDLEWIR RTIPWLEDRS POK  
 !!AA\_SEQUENCE 1 0 PRELIMINARY; PRT; 24 AA.

ID Q9PRV3; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE NAP-185 NEUROGLIA-associated protein (Fragment).  
 OS Gallus gallus (chicken).  
 OC Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archaeosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TAXID=9031;  
 RN [1]  
 RP SQUENCE.  
 RX MEDLINE=95050502; PubMed=7961672;  
 RA Solca F.F., Lurie D.I., Diltz C.D., Johnson R.S., Kumar S.,  
 RA Rubel F.W., Fischer E.H.;  
 RT "Identification and purification of a chicken brain neuroglia-  
 associated protein.";  
 RL J. Biol. Chem. 269:27559-27565(1994).  
 SQ SEQUENCE 24 AA; 2404 MW; 73E0E6067F253442 CRC64;

Q9PRV3 Length: 24 December 22, 2002 19:30 Type: P Check: 2489 ..

1 DGGE DRDAV EAVNLGTGGC RTPK  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 25 AA.

ID Q9DU42; 01-MAR-001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Rev protein (Fragment).  
 GN REV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TAXID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MDDU2;  
 RX MEDLINE=21002573; PubMed=1118069;  
 RA Motomura K., Kusegawa S., Kato K., Nohtomi K., Lwin H.H., Tun K.M.,  
 RA Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;  
 RT "Emergence of new forms of human immunodeficiency virus type 1  
 intersubtype recombinants in central Myanmar.";  
 RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).  
 DR EMBL: AB04389; BAR19213.1; -.  
 DR InterPro: IPR000628; REV\_protein.  
 PFam: PF000424; REV: 1.  
 FT NON\_TER 25 AA; 25 SEQUENCE 25 AA; 2964 MW; 0F92F42A1DDF2758 CRC64;  
 Q9DU42 Length: 25 December 22, 2002 19:30 Type: P Check: 5339 ..

1 MAGRSERDRDE ELIKTVRLIK LLYQS  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 9 AA.

ID Q70140; 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Tat protein (Fragment).  
 GN TAT.  
 OS Human immunodeficiency virus type 1.  
 OX NCBI\_TAXID=11676;

---

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=020;  
 RX MEDLINE=96194694; PubMed=7888189;  
 RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L., McCuchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;  
 RA RT "Genetic variation of HIV type 1 in four World Health Organization-sponsored vaccine evaluation sites: generation of functional envelope (glycoprotein 160) clones representative of sequence subtypes A, B, C, and E. WHO Network for HIV Isolation and Characterization.";  
 RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=020;  
 RX MEDLINE=96190564; PubMed=8627686;  
 RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S., Karlsson G., Sooroski J., Morgado M., Galvao-Castro B., von Briesen H., Beddoes S., Weber J., Sharp P.M., Shaw G.M., Hahn B.H.;  
 RA RT "Molecular cloning and analysis of functional envelope genes from human immunodeficiency virus type 1 sequence subtypes A through G. The WHO and NIAID Networks for HIV Isolation and Characterization.";  
 RL J. Virol. 70:1651-1657(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=020;  
 RA Allen E.E.;  
 RL Submitted (APR-1994) to the EMBL/GenBank/DDJB databases.  
 DR EMBL: U08794; AACB05175.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 9 AA; 1098 MW; 5B76D4OAB1AB01A3 CRC64;

Q70140 Length: 9 December 22, 2002 19:30 Type: P Check: 3335 ..

1 SKTETDRED  
 !!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 23 AA.

ID Q95575; 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein HP0917.  
 GN HP0917.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OX NCBI\_TAXID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=6695 / ATCC 700392;  
 RX MEDLINE=9394467; PubMed=9352185;  
 RA Tomb J.-F., Kerlavage A.R., Clayton R.A., Sutton G.G., Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A., Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S., Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M., Cottrell M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M., Venter J.C.;  
 RA RT "The complete genome sequence of the gastric pathogen Helicobacter pylori.";  
 RL Nature 368:539-547(1997).  
 DR EMBL: AE000601; AA07968.1; -.  
 DR TIGR: HP0917; -.  
 KW Hypothetical protein; complete proteome.  
 SQ SEQUENCE 23 AA; 2717 MW; FC304239DA2FD44 CRC64;

025575 Length: 23 December 22, 2002 19:30 Type: P Check: 575 ..

1 NSPLPLRNP LTOQDRFFQE IIA

!!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 27 AA.  
ID 006283 DT 01-OCT-2000 (TREMBrel. 15, Last sequence update)  
AC 006283; DT 01-OCT-2000 (TREMBrel. 15, Last sequence update)  
DT 01-JUL-1997 (TREMBrel. 04, Created)  
DT 01-JUL-1997 (TREMBrel. 04, Last sequence update)  
DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
DE Hypothetical protein Rv3599c.  
GN RV3599C OR MTCY07H23.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Actinomycetidae; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1] RP SEQUENCE FROM N.A.  
RN STRAIN=H37RV;  
RX MEDLINE=9825987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parhull J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Elgelmie K.J., Gas S., Barry C.E. III, Tekala F.,  
RA Bedcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagele K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborn J., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence.", Nature 393:537-544(1998).  
RL Nature 393:537-544(1998).  
EMBL; 295557; CAB089451; -.  
DR Tuberculist; RV3599c; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 27 AA; 2894 MW; 1746F2A9YB95C6B CRC64;

006283 Length: 27 December 22, 2002 19:30 Type: P Check: 8577 ..

1 MRASSLGTTG PADRIDLATH ERRREV1

!!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 33 AA.  
ID 09KLT25 DT 01-OCT-2000 (TREMBrel. 15, Created)  
DT 01-OCT-2000 (TREMBrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBrel. 19, Last annotation update)  
DE Hypothetical protein VC0735.  
GN VC0735.

OS Vibrio cholerae.

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; vibrio.

NCBI\_TaxID=666;

[1] RP SEQUENCE FROM N.A.  
RN STRAIN=ST TOR NJ16961 / SEROTYPE O1; MEDLINE=20406833; PubMed=10952301;  
RX Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Rodman R.J., Hafft D.H., Hickie E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Elmajayev M.D., Vanathavan J.H., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.; RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae.", Nature 405:477-483(2000).  
RL Nature 405:477-483(2000).  
EMBL; NE00404; AAC96658.1; -.  
DR EMBL; VCA0761; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 32 AA; 3898 MW; E08860BC8C67F83 CRC64;

Q9KL13 length: 32 December 22, 2002 19:30 Type: P Check: 1300 ..

1 MDRURRESA ICSUDRRKQT NLESSDQIKQ WR

!!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 13 AA.  
ID 08X407 DT 01-MAR-2002 (TREMBrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBrel. 20, Last annotation update)  
DE Hypothetical protein z0899.

GN 20899.

OS Escherichia coli 0157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.

OX NCBI\_TaxID=83334;

RN [1] RP SEQUENCE FROM N.A.  
RN STRAIN=H7 / EDL933 / ATCC 700927; MEDLINE=2107435; PubMed=12106551;  
RX Perma N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirrpatrick H.A.,  
RA Posfai G., Hackney J., Klionsky D.J., Boutin A., Shao Y., Miller L.,  
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potanousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J.J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.; RT "Genome sequence of enterohaemorrhagic Escherichia coli 0157:H7.",  
RA Nature 409:529-535(2000).  
DR EMBL; AE005252; AACG5068.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 13 AA; 1575 MW; 512D35A6BBD763 CRC64;

Q8X407 length: 13 December 22, 2002 19:30 Type: P Check: 6940 ..

1 MSTDRKPVML LFH

!!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 35 AA.  
ID Q8VJ09 AC Q8VJ05;  
DT 01-MAR-2002 (TREMBrel. 20, Created)  
DT 01-MAR-2002 (TREMBrel. 20, Last sequence update)  
DE Hypothetical protein MT3536.

GN MT3536.

OS Mycobacterium tuberculosis.

Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;

!!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 32 AA.  
ID 09KLT3 AC 09KL13;

1 MPVGSSIPCS ISTALADRKV VSRYWCFFTL LVL

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; *Mycobacterium*.  
 OX NCBI\_TAXID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / OSHKOSH;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 Peterson J., Desoye R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 Kolonay J.F., Nelson W.C., Utterback T., Salzberg S.L.,  
 Bechtler A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of *Mycobacterium tuberculosis* clinical and  
 laboratory strains";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AE007158; AAK47876.1; ..  
 DR TIGR; M33536; ..  
 KW Hypothetical protein.  
 SQ SEQUENCE 35 AA; 4204 MW; 897851C24908FC03 CRC64;  
 Q8VJ09 Length: 35 December 22, 2002 19:30 Type: P Check: 8356 ..  
 1 MYKLALIDE YDRADRTERG ALIURRENLYS SLTE  
 !!!AA\_SEQUENCE 1.0 PRELIMINARY; PRT; 32 AA.  
 ID Q9HNL3  
 AC  
 DT 01-MAR-2001 (TREMBREL. 16, Created)  
 DT 01-MAR-2001 (TREMBREL. 16, Last annotation update)  
 DE VNG2049C.  
 GN  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 NCBI\_TAXID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20504483; PubMed=11016950;  
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
 Shukla H.D., Lasky S.R., Baigla N.S., Thorsson V., Sbrogna J.,  
 Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
 Leithauer B., Keller K., Cruz R., Danzon M.J., Hough D.W.,  
 Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isemenbarier T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
 Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
 RT "Genome sequence of *Halobacterium* species NRC-1";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 97:12176-12181(2000).  
 DR EMBL; AE005057; AAG20207.1; ..  
 KW Complete proteome.  
 SQ SEQUENCE 32 AA; 3719 MW; DOF9677E650A578A CRC64;  
 Q9HNL3 Length: 32 December 22, 2002 19:30 Type: P Check: 9872 ..  
 1 MVRVPVTAG RAEMLRADR LETTAFAFWRL VD



!!AA SEQUENCE 1.0  
F1;A21630 - cytochrome P450b - mouse (fragment)  
N;Contains: oxidoreductase (EC 1.-.-.-)  
C;Species: Mus musculus (house mouse)  
C;Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 11-Jun-1999  
C;Accession: A21630  
R;Stupans, I.; Ikeda, T.; Kessler, D.J.; Nebert, D.W.  
DNA 3, 129-137, 1984  
A;Title: Characterization of a cDNA clone for mouse phenobarbital-inducible cytochrome p-450b.  
A;Reference number: A21630; MUID:84207435; PMID:6547088  
A;Accession: A21630  
A;Molecule type: mRNA  
A;Residues: 1-25 <STU>  
A;Cross-references: GB:K02409; NID:g192896; PIDN:AAA37510.1; PID:g553905  
A;Note: the authors translated the codon CTA for residue 5 as Pro, ACC for residue 7 as Ser and AGT for residue 12 as Thr  
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C;Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protein

A21630 Length: 25 December 22, 2002 19:21 Type: P Check: 4925 ..

1 SHRLPTLDDR SKMPYTDADI HEIQR

!!AA SEQUENCE 1.0  
F1;F42762 - C 3.4.25.1 proteasome endopeptidase complex () subunit 13 - bovine (fragment)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 03-Jun-2002  
C;Accession: F42762  
R;Dick, L.R.; Moomaw, C.R.; Pramanik, B.C.; DeMartino, G.N.; Slaughter, C.A.  
Biochemistry 31, 7347-7355, 1992  
A;Title: Identification and localization of a cysteinyl residue critical for the trypsin-like catalytic activity of the proteasome.  
A;Reference number: A42762; MUID:92378961; PMID:1510924  
A;Accession: F42762  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <DIC>  
A;Note: sequence extracted from NCBI backbone (NCBIP:112175)  
C;Superfamily: multicatalytic endopeptidase complex chain C9  
C;Keywords: hydrolase

F42762 Length: 20 December 22, 2002 19:21 Type: P Check: 5572 ..

1 MKGKNXVAIA ADRREGIQAQ

!!AA SEQUENCE 1.0  
F1;S55729 - orotidine-5'-monophosphate decarboxylase - Aspergillus awamori (fragments)  
C;Species: Aspergillus awamori  
C;Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
C;Accession: S55729  
R;Gouka, R.J.; Hessing, J.G.M.; Stam, H.; Musters, W.; van den Hondel, C.A.M.J.J.  
Curr. Genet. 27, 536-540, 1995  
A;Title: A novel strategy for the isolation of defined pyrG mutants and the development of a site-specific integration system for Aspergillus awamori.  
A;Reference number: S55729; MUID:96031709; PMID:7553938  
A;Accession: S55729  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-18;19-28 <GOU>  
C;Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate decarboxylase homology

S55729 Length: 28 December 22, 2002 19:21 Type: P Check: 1645 ..

1 VTVSADVTTT KELLDLADRV MGFVSTRS



!!AA\_SEQUENCE 1.0  
P1;B44336 - neurotoxin Tx3-2 - spider (Phoneutria nigriventer)  
C;Species: Phoneutria nigriventer  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Aug-1998  
C;Accession: B44336  
R;Cordeiro M do, N.; de Figueiredo, S.G.; Valentim A do, C.; Diniz, C.R.; von Eickstedt, V.R.; Gilroy, J.; Richardson, M.  
Toxicon 31, 35-42, 1993  
A;Title: Purification and amino acid sequences of six Tx3 type neurotoxins from the venom of the Brazilian 'armed' spider Phoneutria nigriventer (Keys.).  
A;Reference number: A44336; MUID:93190315; PMID:8446961  
A;Accession: B44336  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-34 <COR>  
A;Experimental source: venom  
A;Note: sequence extracted from NCBI backbone (NCBIP:126773)  
C;Superfamily: curtatoxin  
C;Keywords: neurotoxin

B44336 Length: 34 December 22, 2002 19:21 Type: P Check: 2946 ..

1 ACAGLYKKCG KGASPCCEDR PCKCDLAMGN CICK

!!AA\_SEQUENCE 1.0  
F1;B28563 - hemoglobin chain IV - earthworm (Lumbricus terrestris) (fragment)  
C;Species: Lumbricus terrestris (common earthworm)  
C;Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 04-Mar-2000  
C;Accession: B28563  
R;Gotoh, T.; Shishikura, F.; Snow, J.W.; Ereifej, K.I.; Vinogradov, S.N.; Walz, D.A.  
Biochem. J. 241, 441-445, 1987  
A;Title: Two globin strains in the giant annelid extracellular haemoglobins.  
A;Reference number: A90337; MUID:87241210; PMID:3593201  
A;Accession: B28563  
A;Molecule type: protein  
A;Residues: 1-22 <GOT>  
C;Superfamily: globin; globin homology  
C;Keywords: oxygen carrier

B28563 Length: 22 December 22, 2002 19:21 Type: P Check: 9362 ..

1 ADEEESCCSYE DRREVRHIWD DV

!!AA\_SEQUENCE 1.0  
F1;S01808 - hemoglobin AII - tube worm (Lamellibrachia sp.) (fragment)  
C;Species: Lamellibrachia sp.  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 04-Mar-2000  
C;Accession: S01808  
R;Suzuki, T.; Takagi, T.; Ohta, S.  
Biochem. J. 255, 541-545, 1988  
A;Title: N-terminal amino acid sequence of the deep-sea tube worm haemoglobin remarkably resembles that of annelid haemoglobin.  
A;Reference number: S01807; MUID:89076216; PMID:3202832  
A;Accession: S01808  
A;Molecule type: protein  
A;Residues: 1-24 <SUZ>  
C;Superfamily: globin; globin homology  
C;Keywords: oxygen carrier

S01808 Length: 24 December 22, 2002 19:21 Type: P Check: 3406 ..

1 SSNSCTTEDR REMQLMWANV WSAQ

!!AA\_SEQUENCE 1.0  
F1;D60894 - gamma crystallin V - bullfrog (fragment)  
C;Species: Rana catesbeiana (bullfrog)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Nov-1997  
C;Accession: D60894  
R;Chiou, S.H.  
J. Protein Chem. 7, 527-534, 1988



A;Title: The protein sequence homology of gamma-crystallins among major vertebrate classes and their DNA sequence homology to heat-shock protein genes.  
A;Reference number: A60894; MUID:89351593; PMID:3255376  
A;Accession: D60894  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
C;Superfamily: beta-crystallin  
C;Keywords: duplication; eye lens

D60894 Length: 20 December 22, 2002 19:21 Type: P Check: 6311 ..

1 GKIVFYEDRN FQGRSYECSS

!!AA\_SEQUENCE 1.0  
F1;S04621 - gamma-crystallin II - milk shark (fragment)  
C;Species: Rhizoprionodon acutus (milk shark)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Nov-1997  
C;Accession: S04621  
R;Chiou, S.H.  
FEBS Lett. 250, 25-29, 1989  
A;Title: Characterization of gamma-crystallins from eye lenses of shark: closer structural similarity to mammalian than other piscine gamma-crystallins?  
A;Reference number: S04621; MUID:89290026; PMID:2737298  
A;Accession: S04621  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
C;Superfamily: beta-crystallin  
C;Keywords: duplication

S04621 Length: 20 December 22, 2002 19:21 Type: P Check: 5853 ..

1 GKITFYEDRG FQGHCYECSS

!!AA\_SEQUENCE 1.0  
F1;A61570 - gamma-crystallin (total) - siamang (fragment)  
C;Species: Hylobates syndactylus (siamang)  
C;Date: 25-Oct-1994 #sequence\_revision 06-Jan-1995 #text\_change 21-Nov-1997  
C;Accession: A61570  
R;Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.  
Curr. Eye Res. 7, 1017-1022, 1988  
A;Title: Biochemical characterization of lens crystallins from three mammalian species.  
A;Reference number: A61570; MUID:89152306; PMID:3229121  
A;Accession: A61570  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
A;Note: 8-Glu, 9-Gln, 10-Ala, 15-Ser, 18-Thr, and 20-Ser were also found  
A;Note: this preparation appears to represent the product of several gamma crystallin genes from one specimen  
C;Superfamily: beta-crystallin  
C;Keywords: duplication; eye lens

A61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6107 ..

1 GKITFYEDRG FQGRHYECTT

!!AA\_SEQUENCE 1.0  
F1;B61570 - gamma-crystallin (total) - Formosa sika deer (fragment)  
C;Species: Cervus nippon taiouanus (Formosa sika deer)  
C;Date: 25-Oct-1994 #sequence\_revision 06-Jan-1995 #text\_change 21-Nov-1997  
C;Accession: B61570  
R;Chiou, S.H.; Chang, W.P.; Ting, L.M.; Lai, T.A.; Lin, H.K.  
Curr. Eye Res. 7, 1017-1022, 1988  
A;Title: Biochemical characterization of lens crystallins from three mammalian species.  
A;Reference number: A61570; MUID:89152306; PMID:3229121  
A;Accession: B61570  
A;Molecule type: protein  
A;Residues: 1-20 <CHI>  
A;Note: 6-Phe, 14-His, and 15-Cys were also found  
A;Note: this preparation appears to represent the product of several gamma



crystallin genes from one specimen  
C;Superfamily: beta-crystallin  
C;Keywords: duplication; eye lens

B61570 Length: 20 December 22, 2002 19:21 Type: P Check: 6068 ..

1 GKITFYEDRG FQGRHYECSS

!!AA\_SEQUENCE 1.0  
F1;A60502 - myonexin - northern leopard frog (fragment)  
C;Species: Rana pipiens (northern leopard frog)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C;Accession: A60502; A33087  
R;Tidball, J.G.  
Dev. Biol. 142, 103-114, 1990  
A;Title: Myonexin: an 80-kDa glycoprotein that binds fibronectin and is located at embryonic myotendinous junctions.  
A;Reference number: A60502; MUID:91032554; PMID:2146177  
A;Accession: A60502  
A;Molecule type: protein  
A;Residues: 1-25 <TID>  
C;Comment: This protein resembles calsequestrin in sequence but differs in distribution and physical properties.  
C;Superfamily: calsequestrin  
C;Keywords: fibronectin binding; glycoprotein; skeletal muscle

A60502 Length: 25 December 22, 2002 19:21 Type: P Check: 4514 ..

1 EDGFDFPEYD GEDRVIFISL KNYKA

!!AA\_SEQUENCE 1.0  
F1;E33208 - calreticulin, uterine - rabbit (fragment)  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 31-Jul-1991 #sequence\_revision 31-Jul-1991 #text\_change 01-Mar-1996  
C;Accession: E33208; F33208  
R;Milner, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance, J.E.; Opas, M.; Michalak, M.  
J. Biol. Chem. 266, 7155-7165, 1991  
A;Title: Calreticulin, and not calsequestrin, is the major calcium binding protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.  
A;Reference number: A33208; MUID:91201375; PMID:2016321  
A;Accession: E33208  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-29 <MI2>  
A;Note: uterine form  
A;Accession: F33208  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <MI2>  
A;Note: hepatic form  
C;Superfamily: calreticulin

E33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2821 ..

1 EPVVFKEQF LDGDGWTDRW IESKHKSDP

!!AA\_SEQUENCE 1.0  
F1;C33208 - calreticulin, slow twitch skeletal muscle - rabbit (fragment)  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 31-Jul-1991 #sequence\_revision 31-Jul-1991 #text\_change 07-Feb-1997  
C;Accession: C33208  
R;Milner, R.E.; Baksh, S.; Shemanko, C.; Carpenter, M.R.; Smillie, L.; Vance, J.E.; Opas, M.; Michalak, M.  
J. Biol. Chem. 266, 7155-7165, 1991  
A;Title: Calreticulin, and not calsequestrin, is the major calcium binding protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic reticulum.  
A;Reference number: A33208; MUID:91201375; PMID:2016321  
A;Accession: C33208  
A;Status: preliminary  
A;Molecule type: mRNA



A;Residues: 1-29 <MIL>  
C;Superfamily: calreticulin  
C;Keywords: skeletal muscle

C33208 Length: 29 December 22, 2002 19:21 Type: P Check: 2871 ..

1 EPVVFQFQF LDGDGWTDRW IESKHKSDF

!!AA\_SEQUENCE 1.0  
F1;A33434 - calcium-binding protein - dog (fragment)  
C;Species: Canis lupus familiaris (dog)  
C;Date: 21-Feb-1990 #sequence\_revision 21-Feb-1990 #text\_change 16-Feb-1997  
C;Accession: A33434  
R;Collins, J.H.; Xi, Z.; Alderson-Lang, B.H.; Treves, S.; Volpe, P.  
Biochem. Biophys. Res. Commun. 164, 575-579, 1989  
A;Title: Sequence homology of a canine brain calcium-binding protein with  
calregulin and the human Ro/SS-A antigen.  
A;Reference number: A33434; MUID:90026437; PMID:2803321  
A;Accession: A33434  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-24 <COL>  
C;Superfamily: calreticulin  
C;Keywords: calcium binding

A33434 Length: 24 December 22, 2002 19:21 Type: P Check: 2817 ..

1 EPAIYFKEQF LDGDGXTRDX IESK

!!AA\_SEQUENCE 1.0  
F1;B41481 - virulence-associated protein (virA 5' region) - Salmonella  
typhimurium (fragment)  
C;Species: Salmonella typhimurium  
C;Date: 10-Apr-1992 #sequence\_revision 10-Apr-1992 #text\_change 20-Aug-1999  
C;Accession: B41481  
R;Gulig, P.A.; Chiodo, V.A.  
Infect. Immun. 58, 2651-2658, 1990  
A;Title: Genetic and DNA sequence analysis of the Salmonella typhimurium  
virulence plasmid gene encoding the 28,000-molecular-weight protein.  
A;Reference number: A41481; MUID:90316693; PMID:2164511  
A;Accession: B41481  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-36 <GUL>  
A;Cross-references: GB:M64295; GB:M33927; NID:g154440; PIDN:AAA27272.1;  
PID:g154441  
C;Superfamily: virulence-associated protein spvB  
C;Keywords: DNA binding; transcription regulation

B41481 Length: 36 December 22, 2002 19:21 Type: P Check: 1642 ..

1 GSQDFASQLS KLRLSDDRTA DTNRIKRIIN MRVLNS

!!AA\_SEQUENCE 1.0  
F1;S07394 - DNA-invertase - phage D108 (fragment)  
C;Species: phage D108  
A;Note: host Escherichia coli  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Jun-2000  
C;Accession: S07394  
R;Szatmari, G.B.; Lapointe, M.; DuBow, M.S.  
Nucleic Acids Res. 15, 6691-6704, 1987  
A;Title: The right end of transposable bacteriophage D108 contains a 520 base  
pair protein-encoding sequence not present in bacteriophage Mu.  
A;Reference number: S07394; MUID:87316928; PMID:2957646  
A;Accession: S07394  
A;Molecule type: DNA  
A;Residues: 1-22 <SZA>  
A;Cross-references: EMBL:X05926; NID:g14918; PIDN:CAA29365.1; PID:g1364174  
C;Genetics:  
A;Gene: gin  
C;Superfamily: transposase repressor



C;Keywords: DNA binding; DNA integration; DNA recombination

S07394 Length: 22 December 22, 2002 19:21 Type: P Check: 9004 ..

1 YKKHPAKRTH IENDDRINQI DR

!!AA\_SEQUENCE 1.0

F1;S35552 - vesicle-associated membrane protein 1 - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999

C;Accession: S35552

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.

Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35552

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-32 <PAT>

C;Superfamily: synaptobrevin

S35552 Length: 32 December 22, 2002 19:21 Type: P Check: 9008 ..

1 ERDQKLSELD DRADALQAGA SQFESSAAKL KR

!!AA\_SEQUENCE 1.0

F1;S35553 - vesicle-associated membrane protein 1 - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 20-May-1994 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999

C;Accession: S35553

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.

Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35553

A;Molecule type: DNA

A;Residues: 1-32 <PAT>

C;Superfamily: synaptobrevin

S35553 Length: 32 December 22, 2002 19:21 Type: P Check: 9118 ..

1 ERDQKLSELD DRADALQAGA SVFESSAAKL KR

!!AA\_SEQUENCE 1.0

F1;S35555 - vesicle-associated membrane protein 2 - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 10-Dec-1993 #sequence\_revision 26-Jul-1996 #text\_change 17-Mar-1999

C;Accession: S35555

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.

Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916

A;Accession: S35555

A;Molecule type: DNA

A;Residues: 1-32 <PAT>

C;Superfamily: synaptobrevin

C;Keywords: membrane trafficking; transmembrane protein

S35555 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..

1 ERDQKLSELD DRADALQAGA SQFETSAAKL KR

!!AA\_SEQUENCE 1.0

F1;S35554 - vesicle-associated membrane protein 2 - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 17-Mar-1999

C;Accession: S35554

R;Patarnello, T.; Bargelloni, L.; Rossetto, O.; Schiavo, G.; Montecucco, C.

Nature 364, 581-582, 1993

A;Title: Neurotransmission and secretion.

A;Reference number: S35552; MUID:93354436; PMID:8350916



A;Accession: S35554  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-32 <PAT>  
C;Superfamily: synaptobrevin

S35554 Length: 32 December 22, 2002 19:21 Type: P Check: 9033 ..

1 ERDQKLSELD DRADALQAGA SQFETSAAKL KR

!!AA\_SEQUENCE 1.0  
F1;E44621 - homeotic protein Hox 10 (clone 10w) - sea lamprey (fragment)  
C;Species: Petromyzon marinus (sea lamprey)  
C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 15-Oct-1999  
C;Accession: E44621  
R;Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993  
A;Title: Expansion of the Hox gene family and the evolution of chordates.  
A;Reference number: A44616; MUID:93317669; PMID:8101001  
A;Accession: E44621  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-27 <PEN>  
A;Cross-references: GB:L14895; NID:g290813; PIDN:AAA02530.1; PID:g290814  
C;Superfamily: unassigned homeobox proteins; homeobox homology  
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

E44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9552 ..

1 LFNMYLTRER RLEISRGVNL TDRQVKI

!!AA\_SEQUENCE 1.0  
F1;F44621 - homeotic protein Hox 10 (clone 10x) - sea lamprey (fragment)  
C;Species: Petromyzon marinus (sea lamprey)  
C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 15-Oct-1999  
C;Accession: F44621  
R;Pendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993  
A;Title: Expansion of the Hox gene family and the evolution of chordates.  
A;Reference number: A44616; MUID:93317669; PMID:8101001  
A;Accession: F44621  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-27 <PEN>  
A;Cross-references: GB:L14900; NID:g290819; PIDN:AAA02535.1; PID:g290820  
C;Superfamily: unassigned homeobox proteins; homeobox homology  
C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F44621 Length: 27 December 22, 2002 19:21 Type: P Check: 9407 ..

1 LFSMYLTRER RLEISHLLSL TDRQVKI

!!AA\_SEQUENCE 1.0  
P1;C84355 - hypothetical protein Vng2049c [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: C84355  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.;  
Lasky, S.R.; Baliga, N.; Thorsson, V.; Sbrogna, J.; Swartzell, S.; Weir, D.;  
Hall, J.; Dahl, T.A.; Welti, R.; Goo, Y.A.; Leithauser, B.; Keller, K.; Cruz,  
R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonski, P.E.; Krebs, M.P.;  
Angevine, C.M.; Dale, H.; Isenbarger, T.A.; Peck, R.F.; Pohlschrod, M.;  
Spudich, J.L.; Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe,  
T.M.; Liang, P.; Riley, M.; Hood, L.; DasSarma, S.  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: C84355  
A;Status: preliminary  
A;Molecule type: DNA



A;Residues: 1-32 <STO>  
A;Cross-references: GB:AE004437; NID:g10581474; PIDN:AAG20207.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: VNG2049C

C84355 Length: 32 December 22, 2002 19:21 Type: P Check: 9872 ..

1 MVRVPVTDAG RAEMRERADR LETTAAFWRL VD

!!AA\_SEQUENCE 1.0  
P1;E64634 - hypothetical protein HP0917 - *Helicobacter pylori* (strain 26695)  
C;Species: *Helicobacter pylori*  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C;Accession: E64634  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.;  
Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.;  
Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus,  
B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.;  
Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne,  
J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman,  
J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.;  
Fraser, C.M.; Venter, J.C.  
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: E64634  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-23 <TOM>  
A;Cross-references: GB:AE000601; GB:AE000511; NID:g2314051; PIDN:AAD07968.1;  
PID:g2314059; TIGR:HP0917

E64634 Length: 23 December 22, 2002 19:21 Type: P Check: 575 ..

1 MSPLTPLRNP LTQEDRFFQE IIA

!!AA\_SEQUENCE 1.0  
F1;PH0858 - MauD protein - *Paracoccus denitrificans* (fragment)  
C;Species: *Paracoccus denitrificans*  
C;Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Oct-1999  
C;Accession: PH0858  
R;Chistoserdov, A.Y.; Boyd, J.; Mathews, F.S.; Lidstrom, M.E.  
Biochem. Biophys. Res. Commun. 184, 1181-1189, 1992  
A;Title: The genetic organization of the mau gene cluster of the facultative autotroph *Paracoccus denitrificans*.  
A;Reference number: PH0856; MUID:92272706; PMID:1590782  
A;Accession: PH0858  
A;Molecule type: DNA  
A;Residues: 1-23 <CHI>  
A;Cross-references: GB:M90098; NID:g150580; PIDN:AAA25577.1; PID:g150581  
C;Genetics:  
A;Gene: maud

PH0858 Length: 23 December 22, 2002 19:21 Type: P Check: 807 ..

1 LEADRTGFAS LQQYMASRKK QAA

!!AA\_SEQUENCE 1.0  
F1;A61412 - methane monooxygenase (EC 1.14.13.25) protein Y - *Methylococcus capsulatus* (fragment)  
N;Alternate names: methane hydroxylase protein Y  
C;Species: *Methylococcus capsulatus*  
C;Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Sep-1994  
C;Accession: A61412  
R;Rosenzweig, A.C.; Lippard, S.J.  
Acc. Chem. Res. 27, 229-236, 1994  
A;Title: Determining the structure of a hydroxylase enzyme that catalyzes the conversion of methane to methanol in methanotrophic bacteria.  
A;Reference number: A61412



A;Accession: A61412  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-27 <ROS>  
C;Keywords: oxidoreductase

A61412 Length: 27 December 22, 2002 19:21 Type: P Check: 8072 ..

1 WIEDYASRID FKADRQIVK AVLAGLK

!!AA\_SEQUENCE 1.0  
P1;A82288 - hypothetical protein VC0735 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C;Accession: A82288  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: A82288  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-33 <HEI>  
A;Cross-references: GB:AE004159; GB:AE003852; NID:g9655167; PIDN:AAF93900.1; GSPDB:GN00126; TIGR:VC0735  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC0735  
A;Map position: 1

A82288 Length: 33 December 22, 2002 19:21 Type: P Check: 3787 ..

1 MPVGSSIPCS ISTALADRKV VSRYWCFFTL LVL

!!AA\_SEQUENCE 1.0  
P1;B82421 - hypothetical protein VCA0761 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C;Accession: B82421  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleishmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: B82421  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-32 <HEI>  
A;Cross-references: GB:AE004404; GB:AE003853; NID:g9658174; PIDN:AAF96659.1; GSPDB:GN00127; TIGR:VCA0761  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VCA0761  
A;Map position: 2

B82421 Length: 32 December 22, 2002 19:21 Type: P Check: 1300 ..

1 MPDRLLRFSA ICSTDRRKQT NLFSSDLKQQ WR



!!AA\_SEQUENCE 1.0  
F1;B39089 - hydrogenase (EC 1.18.99.1) 34K chain - Thiocapsa roseopersicina (fragment)  
C;Species: Thiocapsa roseopersicina  
C;Date: 27-Nov-1991 #sequence\_revision 27-Nov-1991 #text\_change 23-Jun-1993  
C;Accession: B39089  
R;Kovacs, K.L.; Tigyi, G.; Thanh, L.T.; Lakatos, S.; Kiss, Z.; Bagyinka, C.  
J. Biol. Chem. 266, 947-951, 1991  
A;Title: Structural rearrangements in active and inactive forms of hydrogenase from Thiocapsa roseopersicina.  
A;Reference number: A39089; MUID:91093297; PMID:1845998  
A;Accession: B39089  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-20 <KOV>  
C;Keywords: oxidoreductase

B39089 Length: 20 December 22, 2002 19:21 Type: P Check: 6231 ..

1 XXEQADRPSV IGLPFQEXTG

!!AA\_SEQUENCE 1.0  
P1;H70954 - hypothetical protein Rv3599c - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C;Accession: H70954  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: H70954  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-27 <COL>  
A;Cross-references: GB:Z95557; GB:AL123456; NID:g3242276; PIDN:CAB08945.1; PID:e316847; PID:g2113977  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: Rv3599c

H70954 Length: 27 December 22, 2002 19:21 Type: P Check: 8577 ..

1 MPASSLGTGS PAADRLDATH ERRREVI

!!AA\_SEQUENCE 1.0  
F1;S14161 - probable dTDP-4-keto-6-deoxy-hexose 3,4-isomerase (EC 5.3.99.-) - Saccharopolyspora erythraea (fragment)  
N;Alternate names: 51K protein  
C;Species: Saccharopolyspora erythraea  
C;Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 21-Nov-1998  
C;Accession: S14161  
R;Caffrey, P.; Green, B.; Packman, L.C.; Rawlings, B.J.; Staunton, J.; Leadlay, P.F.  
Eur. J. Biochem. 195, 823-830, 1991  
A;Title: An acyl-carrier-protein - thioesterase domain from the 6-deoxyerythronolide B synthase of Saccharopolyspora erythraea. High-level production, purification and characterisation in Escherichia coli.  
A;Reference number: S14091; MUID:91153324; PMID:1999200  
A;Accession: S14161  
A;Molecule type: protein  
A;Residues: 1-20 <CAF>  
C;Genetics:



A;Gene: eryCII  
C;Function:  
A;Description: involved in desosamine biosynthesis  
C;Keywords: antibiotic biosynthesis; intramolecular oxidoreductase; isomerase

S14161 Length: 20 December 22, 2002 19:21 Type: P Check: 5957 ..

1 MTTTDRAGLG RQLQMRGLH

!!AA\_SEQUENCE 1.0  
F1;S66213 - glucose 1-dehydrogenase (EC 1.1.1.47) - Haloferax mediterranei  
(fragment)  
C;Species: Haloferax mediterranei  
C;Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 30-Oct-1998  
C;Accession: S66213  
R;Bonete, M.J.; Pire, C.; LLorca, F.I.; Camacho, M.L.  
FEBS Lett. 383, 227-229, 1996  
A;Title: Glucose dehydrogenase from the halophilic Archaeon Haloferax  
mediterranei: enzyme purification, characterisation and N-terminal sequence.  
A;Reference number: S66213; MUID:96198607; PMID:8925901  
A;Accession: S66213  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-17 <BON>  
C;Keywords: oxidoreductase

S66213 Length: 17 December 22, 2002 19:21 Type: P Check: 1672 ..

1 MKAIAVKRGE DRPVVIE

!!AA\_SEQUENCE 1.0  
P1;S23180 - carboxypeptidase - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
C;Accession: S23180  
R;Colombo, S.; D'Auria, S.; Fusi, P.; Zecca, L.; Raia, C.A.; Tortora, P.  
Eur. J. Biochem. 206, 349-357, 1992  
A;Title: Purification and characterization of a thermostable carboxypeptidase  
from the extreme thermophilic archaebacterium Sulfolobus solfataricus.  
A;Reference number: S23180; MUID:92283259; PMID:1597179  
A;Accession: S23180  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-31 <COL>

S23180 Length: 31 December 22, 2002 19:21 Type: P Check: 7956 ..

1 MDLVEKLKND VKEIEDRIIQ IRRKIXENPP L

!!AA\_SEQUENCE 1.0  
P1;S29326 - hypothetical protein 22, psbA 5'-region - Japanese black pine  
chloroplast  
C;Species: chloroplast Pinus thunbergiana (Japanese black pine)  
C;Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 18-Aug-2000  
C;Accession: S29326; T07427  
R;Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Hiratsuka, J.; Shibata, M.;  
Wakasugi, T.; Sugiura, M.  
Mol. Gen. Genet. 232, 206-214, 1992  
A;Title: Chloroplast DNA of black pine retains a residual inverted repeat  
lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and trnH and  
the absence of rps16.  
A;Reference number: S20449; MUID:92212283; PMID:1557027  
A;Accession: S29326  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-22 <TSU>  
A;Cross-references: EMBL:D11467; NID:g344007; PIDN:BAA02023.1; PID:g344011  
R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994  
A;Title: Loss of all ndh genes as determined by sequencing the entire  
chloroplast genome of the black pine *Pinus thunbergii*.



A;Reference number: Z16030; MUID:95024047; PMID:7937893  
A;Accession: T07427  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-22 <WAK>  
A;Cross-references: EMBL:D17510; NID:g529643; PIDN:BAA04307.1; PID:g1262588  
C;Genetics:  
A;Genome: chloroplast  
C;Keywords: chloroplast

S29326 Length: 22 December 22, 2002 19:21 Type: P Check: 9119 ..

1 MEYLTTDRSI ECGIYLLKKIE SI

!!AA\_SEQUENCE 1.0  
F1;T10123 - probable catalase (EC 1.11.1.6) - cucumber (fragment)  
C;Species: Cucumis sativus (cucumber)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C;Accession: T10123  
R;Toyama, T.; Teramoto, H.; Takeba, G.; Tsuji, H.  
Plant Cell Physiol. 36, 1349-1359, 1995  
A;Title: Cytokinin induces a rapid decrease in the levels of mRNAs for catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and other unidentified proteins in etiolated cotyledons of cucumber.  
A;Reference number: Z16946; MUID:96104306; PMID:8564304  
A;Accession: T10123  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-23 <TOY>  
A;Cross-references: EMBL:D63385; NID:g1199475; PIDN:BAA09701.1; PID:g1199476  
A;Experimental source: seedling; cotyledons  
C;Keywords: oxidoreductase

T10123 Length: 23 December 22, 2002 19:21 Type: P Check: 1444 ..

1 NGSQADRSGV QKLAPHLNVR PSI

!!AA\_SEQUENCE 1.0  
F1;PS0212 - 29K protein 4228 - rice (strain Nihonbare) (fragment)  
C;Species: Oryza sativa (rice)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995  
C;Accession: PS0212  
R;Tsugita, A.; Miyatake, N.  
submitted to JIPID, April 1993  
A;Reference number: PS0208  
A;Accession: PS0212  
A;Molecule type: protein  
A;Residues: 1-15 <TSU>  
A;Experimental source: germ  
C;Comment: molecular weight 29K, pI 6.1.

PS0212 Length: 15 December 22, 2002 19:21 Type: P Check: 8963 ..

1 SPADDERRDVG DRYAD

!!AA\_SEQUENCE 1.0  
F1;B60698 - trichocyst protein 27 - Paramecium tetraurelia (fragment)  
C;Species: Paramecium tetraurelia  
C;Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 07-Dec-1999  
C;Accession: B60698  
R;Tindall, S.H.; Devito, L.D.; Nelson, D.L.  
J. Cell Sci. 92, 441-447, 1989  
A;Title: Biochemical characterization of the proteins of Paramecium secretory granules.  
A;Reference number: A60698; MUID:90078398; PMID:2592449  
A;Accession: B60698  
A;Molecule type: protein  
A;Residues: 1-29 <TIN>  
C;Comment: This protein was purified from the extruded matrix (contents) of trichocysts (secretory granules).  
C;Genetics:



A;Genetic code: SGC5  
C;Keywords: extracellular protein

B60698 Length: 29 December 22, 2002 19:21 Type: P Check: 2526 ..

1 DPLDRLLSTL TDLEDRYVAE QKEDDAKNQ

!!AA\_SEQUENCE 1.0  
F1;S19614 - globin - polychaete (*Eudistylia vancouveri*) (fragment)  
N;Alternate names: chlorocruorin  
C;Species: *Eudistylia vancouveri*  
C;Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 05-Dec-1998  
C;Accession: S19614  
R;Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.;  
Kapp, O.H.; Vinogradov, S.N.  
J. Mol. Biol. 222, 1109-1129, 1991  
A;Title: Hierarchy of globin complexes. The quaternary structure of the  
extracellular chlorocruorin of *Eudistylia vancouveri*.  
A;Reference number: S19532; MUID:92106333; PMID:1762147  
A;Accession: S19614  
A;Molecule type: protein  
A;Residues: 1-17 <QAB>  
A;Experimental source: plume  
C;Complex: dodecamers, each consisting of a trimer of tetramers of globin  
chains; dodecamers are linked into bilayer structure by Ca(2+) and  
heme-deficient chimeric globin chains  
C;Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

S19614 Length: 17 December 22, 2002 19:21 Type: P Check: 1820 ..

1 GMKXXSMEDR KTVLADW

!!AA\_SEQUENCE 1.0  
F1;A60741 - insulin-like growth factor-binding protein He39L - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 30-Sep-1993  
C;Accession: A60741  
R;Forbes, B.; Ballard, F.J.; Wallace, J.C.  
J. Endocrinol. 126, 497-506, 1990  
A;Title: An insulin-like growth factor-binding protein purified from medium  
conditioned by a human lung fibroblast cell line (He[39]L) has a novel  
N-terminal sequence.  
A;Reference number: A60741; MUID:91011238; PMID:1698907  
A;Accession: A60741  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-25 <FOR>

A60741 Length: 25 December 22, 2002 19:21 Type: P Check: 4352 ..

1 LAPGPGQGVQ AGXPGGXVEE EDRGG

!!AA\_SEQUENCE 1.0  
F1;B28457 - proteoglycan II, bone - human (tentative sequence) (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 19-May-1989 #sequence\_revision 19-May-1989 #text\_change 31-Mar-2000  
C;Accession: B28457  
R;Fisher, L.W.; Hawkins, G.R.; Tuross, N.; Termine, J.D.  
J. Biol. Chem. 262, 9702-9708, 1987  
A;Title: Purification and partial characterization of small proteoglycans I and  
II, bone sialoproteins I and II, and osteonectin from the mineral compartment  
of developing human bone.  
A;Reference number: A92656; MUID:87250639; PMID:3597437  
A;Accession: B28457  
A;Molecule type: protein  
A;Residues: 1-19 <FIS>

B28457 Length: 19 December 22, 2002 19:21 Type: P Check: 4360 ..

1 DEAXGIAPEV PDDRPFEPS

卷之三

!!AA\_SEQUENCE 1.0  
F1;C54037 - splicing regulatory protein SWAP homolog (alternatively spliced,  
clone pFL2) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 28-Feb-1997  
C;Accession: C54037  
R;Denhez, F.; Lafyatis, R.  
J. Biol. Chem. 269, 16170-16179, 1994  
A;Title: Conservation of regulated alternative splicing and identification of  
functional domains in vertebrate homologs to the Drosophila splicing regulator,  
suppressor-of-white-apricot.  
A;Reference number: A54037; MUID:94266805; PMID:8206918  
A;Accession: C54037  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-29 <DEN>  
A;Experimental source: liver  
A;Note: sequence extracted from NCBI backbone (NCBIP:149928)

C54037 Length: 29 December 22, 2002 19:21 Type: P Check: 1565 ..

1 PQLTQEELEA KQAKQKLEDR LAAAAREKL

!!AA\_SEQUENCE 1.0  
F1;A48845 - sterol regulatory element 1 binding protein (alternatively spliced,  
clone pCY22) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C;Accession: A48845  
R;Yokoyama, C.; Wang, X.; Briggs, M.R.; Admon, A.; Wu, J.; Hua, X.; Goldstein,  
J.L.; Brown, M.S.  
Cell 75, 187-197, 1993  
A;Title: SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls  
transcription of the low density lipoprotein receptor gene.  
A;Reference number: A48845; MUID:94006541; PMID:8402897  
A;Accession: A48845  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-23 <YOK>  
A;Cross-references: GB:S66168; NID:g432596; PIDN:AAB28522.1; PID:g432597  
A;Experimental source: HeLa S3 cells  
A;Note: sequence extracted from NCBI backbone (NCBIN:138566, NCBIN:138568,  
NCBIP:138570)

A48845 Length: 23 December 22, 2002 19:21 Type: P Check: 9975 ..

1 RGRANGTDAP RAGADRGAMD CTF

!!AA\_SEQUENCE 1.0  
F1;S22228 - vitronectin - dog (fragment)  
C;Species: Canis lupus familiaris (dog)  
C;Date: 22-Nov-1993 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997  
C;Accession: S22228  
R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori,  
H.; Matsumoto, I.; Seno, N.; Hayashi, M.  
Biochim. Biophys. Acta 1120, 1-10, 1992  
A;Title: Vitronectin diversity in evolution but uniformity in ligand binding  
and size of the core polypeptide.  
A;Reference number: S21768; MUID:92207982; PMID:1372829  
A;Accession: S22228  
A;Molecule type: protein  
A;Residues: 1-25 <NAK>

S22228 Length: 25 December 22, 2002 19:21 Type: P Check: 4810 ..

1 AQESXKGRVT EGFNADRQQ QDELX

!!AA\_SEQUENCE 1.0  
F1;A26393 - annexin 36K chain - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 18-Jun-1993



C;Accession: A26393  
R;Geisow, M.J.; Fritzsche, U.; Hexham, J.M.; Dash, B.; Johnson, T.  
Nature 320, 636-638, 1986  
A;Title: A consensus amino-acid sequence repeat in Torpedo and mammalian  
Ca(2)+-dependent membrane-binding proteins.  
A;Reference number: A93379; MUID:86203621; PMID:2422556  
A;Accession: A26393  
A;Molecule type: protein  
A;Residues: 1-16 <GEI>

A26393 Length: 16 December 22, 2002 19:21 Type: P Check: 7 ..

1 MLGLGTDEDRLIEIIL

!!AA SEQUENCE 1.0  
P1;S68260 - hypothetical protein gadd7.1 - long-tailed hamster  
C;Species: Cricetulus longicaudatus (long-tailed hamster)  
C;Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C;Accession: S68260  
R;Hollander, M.C.; Alamo, I.; Fornace Jr., A.J.  
Nucleic Acids Res. 24, 1589-1593, 1996  
A;Title: A novel DNA damage-inducible transcript, gadd7, inhibits cell growth,  
but lacks a protein product.  
A;Reference number: S68260; MUID:96211359; PMID:8649973  
A;Accession: S68260  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-38 <HOL>  
A;Cross-references: EMBL:L40430

S68260 Length: 38 December 22, 2002 19:21 Type: P Check: 6375 ..

1 MTPSFSGSSK QLQRNAQMED RGPNHPSEFI TGDNLLLKT

!!AA SEQUENCE 1.0  
P1;A35678 - hypothetical protein (proenkephalin 5' region) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 05-Nov-1999  
C;Accession: A35678  
R;Kilpatrick, D.L.; Zinn, S.A.; Fitzgerald, M.; Higuchi, H.; Sabol, S.L.;  
Meyerhardt, J.  
Mol. Cell. Biol. 10, 3717-3726, 1990  
A;Title: Transcription of the rat and mouse proenkephalin genes is initiated at  
distinct sites in spermatogenic and somatic cells.  
A;Reference number: A35678; MUID:90287163; PMID:2355920  
A;Accession: A35678  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-18 <KIL>  
A;Cross-references: GB:M55181; NID:g201032; PIDN:AAA40127.1; PID:g201033

A35678 Length: 18 December 22, 2002 19:21 Type: P Check: 3193 ..

1 MSSGKQDSPW EDRIPPGR

!!AA SEQUENCE 1.0  
F1;PH0780 - T-cell receptor alpha chain (C11) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PH0780  
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A;Title: T cell receptor genes in a series of class I major histocompatibility  
complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium  
berghei nonapeptide: implications for T cell allelic exclusion and  
antigen-specific repertoire.  
A;Reference number: PH0746; MUID:92078846; PMID:1836010  
A;Accession: PH0780  
A;Molecule type: mRNA  
A;Residues: 1-15 <CAS>  
A;Cross-references: EMBL:X60879



A;Experimental source: T lymphocyte

C;Keywords: T-cell receptor

PH0780 Length: 15 December 22, 2002 19:21 Type: P Check: 8887 ..

1 CALSETGGAD RLTFG

!!AA\_SEQUENCE 1.0

F1;PH0794 - T-cell receptor alpha chain (K1 V-alpha-3.pHDS58) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PH0794

R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.

A;Reference number: PH0746; MUID:92078846; PMID:1836010

A;Accession: PH0794

A;Molecule type: mRNA

A;Residues: 1-17 <CAS>

A;Cross-references: EMBL:X60899

A;Experimental source: T lymphocyte

C;Keywords: T-cell receptor

PH0794 Length: 17 December 22, 2002 19:21 Type: P Check: 1531 ..

1 CAVSMNEYRG ADRLTFG

!!AA\_SEQUENCE 1.0

F1;PH0806 - T-cell receptor alpha chain (PE5.1.1V-alpha-8.F3.3) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PH0806; PH0781

R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.

A;Reference number: PH0746; MUID:92078846; PMID:1836010

A;Accession: PH0806

A;Molecule type: mRNA

A;Residues: 1-15 <CA1>

A;Cross-references: EMBL:X60915

A;Experimental source: T lymphocyte

A;Accession: PH0781

A;Molecule type: mRNA

A;Residues: 1-15 <CA2>

A;Cross-references: EMBL:X60880

A;Experimental source: T lymphocyte

C;Keywords: T-cell receptor

PH0806 Length: 15 December 22, 2002 19:21 Type: P Check: 8864 ..

1 CALSDQGGAD RLTFG

!!AA\_SEQUENCE 1.0

F1;S03505 - T-cell receptor alpha chain J region (80) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 30-May-1997

C;Accession: S03505

R;Winoto, A.; Mjolsness, S.; Hood, L.

Nature 316, 832-836, 1985

A;Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.

A;Reference number: S03503; MUID:85296332; PMID:2993908

A;Accession: S03505

A;Molecule type: DNA



A;Residues: 1-20 <WIN>  
A;Cross-references: EMBL:X02859  
A;Note: this sequence was determined from the germline gene  
C;Keywords: T-cell receptor

S03505 Length: 20 December 22, 2002 19:21 Type: P Check: 6013 ..

1 NTEGADRLTF GKGTQLIIQP

!!AA\_SEQUENCE 1.0  
F1;PT0212 - T-cell receptor alpha chain V-J region (4-1-E.2) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 30-May-1997  
C;Accession: PT0212  
R;Nakano, N.; Kikutani, H.; Nishimoto, H.; Kishimoto, T.  
J. Exp. Med. 173, 1091-1097, 1991  
A;Title: T cell receptor V gene usage of islet beta cell-reactive T cells is not restricted in non-obese diabetic mice.  
A;Reference number: PT0209; MUID:91217621; PMID:1902501  
A;Accession: PT0212  
A;Molecule type: mRNA  
A;Residues: 1-10 <NAK>  
C;Keywords: T-cell receptor

PT0212 Length: 10 December 22, 2002 19:21 Type: P Check: 3993 ..

1 CAVAGGADRL

!!AA\_SEQUENCE 1.0  
F1;PT0547 - T-cell receptor beta chain V-D-J region (126-1AI) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0547  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0547  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-8 <FEE>  
A;Experimental source: day 18 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

PT0547 Length: 8 December 22, 2002 19:21 Type: P Check: 2627 ..

1 ASSDADRG

!!AA\_SEQUENCE 1.0  
F1;PT0676 - T-cell receptor beta chain V-D-J region (140-1AL) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0676  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0676  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-7 <FEE>  
A;Experimental source: day 18 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

PT0676 Length: 7 December 22, 2002 19:21 Type: P Check: 2049 ..

1 ASGEDRG

!!AA\_SEQUENCE 1.0



F1;PT0576 - T-cell receptor beta chain V-D-J region (141-1G) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PT0576  
R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0576  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-7 <FEE>  
A;Experimental source: day 19 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

PT0576 Length: 7 December 22, 2002 19:21 Type: P Check: 2172 ..

1 ASSDDRT

!!AA\_SEQUENCE 1.0  
F1;PT0366 - T-cell receptor beta chain V-J region (6R2) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 30-May-1997  
C;Accession: PT0366  
R;Lehmann, P.V.; Drexler, K.; Tary-Lehmann, M.; Falcioni, F.; Hurtenbach, U.;  
Nagy, Z.A.  
J. Exp. Med. 173, 333-341, 1991  
A;Title: Graft-versus-host resistance induced by class II major histocompatibility complex-specific T cell clones.  
A;Reference number: PT0360; MUID:91108330; PMID:1824856  
A;Accession: PT0366  
A;Molecule type: mRNA  
A;Residues: 1-28 <LEH>  
C;Keywords: T-cell receptor

PT0366 Length: 28 December 22, 2002 19:21 Type: P Check: 480 ..

1 LYFCASSEDR NNQLRFLERG LDHSVLED

!!AA\_SEQUENCE 1.0  
F1;A46592 - lactase-phlorizin hydrolase, 200K isoform - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 01-Nov-1996  
C;Accession: A46592  
R;Dudley, M.A.; Hachey, D.L.; Quaroni, A.; Hutchens, T.W.; Nichols, B.L.;  
Rosenthaler, J.; Perkinson, J.S.; Cook, G.; Reeds, P.J.  
J. Biol. Chem. 268, 13609-13616, 1993  
A;Title: In vivo sucrase-isomaltase and lactase-phlorizin hydrolase turnover in the fed adult rat.  
A;Reference number: A46592; MUID:93293888; PMID:8514793  
A;Accession: A46592  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-17 <DUD>  
A;Note: sequence extracted from NCBI backbone (NCBIP:134559)  
C;Keywords: carbohydrate digestion; intestine

A46592 Length: 17 December 22, 2002 19:21 Type: P Check: 1363 ..

1 DWEDRNFIAA GPLTNLD

!!AA\_SEQUENCE 1.0  
F1;S78414 - ribosomal protein RL25, mitochondrial [validated] - rat (tentative sequence) (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jul-2000  
C;Accession: S78414  
R;Goldschmidt-Reisin, S.; Graack, H.R.  
submitted to the Protein Sequence Database, February 1998  
A;Reference number: S78411



A;Accession: S78414  
A;Molecule type: protein  
A;Residues: 1-26 <GOL>  
A;Note: the protein is designated as mitochondrial ribosomal protein L25  
C;Keywords: mitochondrion; protein biosynthesis; ribosome

S78414 Length: 26 December 22, 2002 19:21 Type: P Check: 6086 ..

1 ISRRXEKKNK IVYPDQLDGE DRRDAE

!!AA\_SEQUENCE 1.0  
F1;S22227 - vitronectin - guinea pig (fragment)  
C;Species: Cavia porcellus (guinea pig)  
C;Date: 22-Nov-1993 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997  
C;Accession: S22227  
R;Nakashima, N.; Miyazaki, K.; Ishikawa, M.; Yatohgo, T.; Ogawa, H.; Uchibori, H.; Matsumoto, I.; Seno, N.; Hayashi, M.  
Biochim. Biophys. Acta 1120, 1-10, 1992  
A;Title: Vitronectin diversity in evolution but uniformity in ligand binding and size of the core polypeptide.  
A;Reference number: S21768; MUID:92207982; PMID:1372829  
A;Accession: S22227  
A;Molecule type: protein  
A;Residues: 1-25 <NAK>

S22227 Length: 25 December 22, 2002 19:21 Type: P Check: 5585 ..

1 XXESXKGRT EGFnADRkXQ XXELX

!!AA\_SEQUENCE 1.0  
F1;A54226 - light-harvesting protein B-830 alpha-1 chain - Chromatium purpuratum (fragment)  
C;Species: Chromatium purpuratum  
C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-Jan-1996  
C;Accession: A54226  
R;Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.  
Biochemistry 33, 2178-2184, 1994  
A;Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.  
A;Reference number: A54226; MUID:94162224; PMID:8117674  
A;Accession: A54226  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-26 <KER>  
C;Keywords: antenna complex; light-harvesting polypeptide

A54226 Length: 26 December 22, 2002 19:21 Type: P Check: 6170 ..

1 MKVPVMMADE NAKLNNPEDD RKKFFV

!!AA\_SEQUENCE 1.0  
F1;B54226 - light-harvesting protein B-830 alpha-2 chain - Chromatium purpuratum (fragment)  
C;Species: Chromatium purpuratum  
C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-Jan-1996  
C;Accession: B54226  
R;Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.  
Biochemistry 33, 2178-2184, 1994  
A;Title: Purification and characterization of the peripheral antenna of the purple-sulfur bacterium Chromatium purpuratum: evidence of an unusual pigment-protein composition.  
A;Reference number: A54226; MUID:94162224; PMID:8117674  
A;Accession: B54226  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-23 <KER>  
C;Keywords: antenna complex; light-harvesting polypeptide

B54226 Length: 23 December 22, 2002 19:21 Type: P Check: 634 ..



1 MQVPVMLADK NAKLNHPEDD RKR

!!AA\_SEQUENCE 1.0  
F1;T46593 - phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)  
C;Species: Mycobacterium marinum  
C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
C;Accession: T46593  
R;Ramakrishnan, L.; Tran, H.T.; Federspiel, N.A.; Falkow, S.  
J. Bacteriol. 179, 5862-5868, 1997  
A;Title: A crtB homolog essential for photochromogenicity in Mycobacterium marinum: isolation, characterization, and gene disruption via homologous recombination.  
A;Reference number: Z23096; MUID:97440138; PMID:9294446  
A;Accession: T46593  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-38 <RAM>  
A;Cross-references: EMBL:U92075; NID:g1928930; PIDN:AAB71427.1; PID:g1928931  
A;Experimental source: strain M  
C;Genetics:  
A;Gene: crtI

T46593 Length: 38 December 22, 2002 19:21 Type: P Check: 6757 ..

1 VPGVGVPPTL ISGRLAADRI TGNTTRSIRH LDLKAQLS

!!AA\_SEQUENCE 1.0  
P1;H85575 - hypothetical protein Z0899 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: H85575  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: H85575  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-13 <STO>  
A;Cross-references: GB:AE005174; NID:g12513665; PIDN:AAG55068.1; GSPDB:GN00145; UWGP:Z0899  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: Z0899

H85575 Length: 13 December 22, 2002 19:21 Type: P Check: 6940 ..

1 MSTDRKPVML LFH

!!AA\_SEQUENCE 1.0  
F1;PA0041 - plastoquinol-plastocyanin reductase (EC 1.10.99.1) - Arabidopsis thaliana (fragment)  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 03-Jun-2002  
C;Accession: PA0041  
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
submitted to JIPID, July 1994  
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensional gel electrophoresis.  
A;Reference number: PA0001  
A;Accession: PA0041  
A;Molecule type: protein  
A;Residues: 1-15 <KAM>  
A;Experimental source: leaf  
C;Keywords: oxidoreductase



PA0041 Length: 15 December 22, 2002 19:21 Type: P Check: 9117 ..

1 ASSIPADRV DMEKR



!!AA\_SEQUENCE 1.0  
ID AL20\_CARMA STANDARD; PRT; 27 AA.  
AC P81823;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Carcinustatin 20.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
OC Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
RA Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the  
RT allatostatin superfamily in the shore crab Carcinus maenas.";  
RL Eur. J. Biochem. 250:727-734(1997).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation; Multigene family.  
FT MOD\_RES 27 27 AMIDATION (POTENTIAL).  
SQ SEQUENCE 27 AA; 3152 MW; 597A6901965FE987 CRC64;

AL20\_CARMA Length: 27 December 22, 2002 19:30 Type: P Check: 9046 ..

1 GYEDEDEDRP FYALGLGKRP RTYSFGL

!!AA\_SEQUENCE 1.0  
ID AMD1\_CHICK STANDARD; PRT; 26 AA.  
AC P81073;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase  
DE isoform M) (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Breast;  
RX MEDLINE=97269365; PubMed=9114497;  
RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.;  
RT "AMP-deaminases from chicken and rabbit muscle: partial primary  
RT sequences of homologous 17-kDa CNBr fragments: autorecognition by  
RT rabbit anti-[chicken AMPD].";  
RL Comp. Biochem. Physiol. 116B:371-377(1997).  
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY  
CC METABOLISM.  
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.  
DR InterPro; IPRO001365; A/AMP\_deaminase.  
DR PROSITE; PS00485; A\_DEAMINASE; PARTIAL.  
KW Hydrolase; Nucleotide metabolism; Multigene family.  
FT NON\_TER 1 1  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 3195 MW; B03E296D63BB6E75 CRC64;

AMD1\_CHICK Length: 26 December 22, 2002 19:30 Type: P Check: 7006 ..

1 MNQKHLLRFI KKSYRVDADR VVYDAK

!!AA\_SEQUENCE 1.0



ID AMD1\_RABIT STANDARD; PRT; 26 AA.  
AC P81072;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isoform M) (Fragment).  
GN AMPD1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=97269365; PubMed=9114497;  
RA Chilson O.P., Kelly-Chilson A.E., Siegel N.R.;  
RT "AMP-deaminases from chicken and rabbit muscle: partial primary  
RT sequences of homologous 17-kDa CNBr fragments: autorecognition by  
RT rabbit anti-[chicken AMPD].";  
RL Comp. Biochem. Physiol. 116B:371-377(1997).  
CC -!- FUNCTION: AMP DEAMINASE PLAYS 'A CRITICAL ROLE IN ENERGY  
CC METABOLISM.  
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP  
CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP  
CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,  
CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3  
CC IS FOUND IN ERYTHROCYTES.  
CC -!- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.  
DR InterPro; IPR001365; A/AMP\_deaminase.  
DR PROSITE; PS00485; A\_DEAMINASE; PARTIAL.  
KW Hydrolase; Nucleotide metabolism; Multigene family.  
FT NON\_TER 1 1  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 3169 MW; B022467EACBB6E75 CRC64;

AMD1\_RABIT Length: 26 December 22, 2002 19:30 Type: P Check: 7826 ..

1 MNQKHLLRFI KKSYQVDADR VVYSTK

!!AA SEQUENCE 1.0  
ID DNIV\_BPD10 STANDARD; PRT; 22 AA.  
AC Q38199;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE DNA-invertase (Fragment).  
GN GIN.  
OS Bacteriophage D108.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC Mu-like viruses.  
OX NCBI\_TaxID=10671;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87316928; PubMed=2957646;  
RA Szatmari G.B., Lapointe M., Dubow M.S.;  
RT "The right end of transposable bacteriophage D108 contains a 520 base  
RT pair protein-encoding sequence not present in bacteriophage Mu.";  
RL Nucleic Acids Res. 15:6691-6703(1987).  
CC -!- FUNCTION: THIS PROTEIN CATALYZES THE INVERSION OF A 3000-BP  
CC SEGMENT OF PHAGE DNA. THE INVERSION RESULTS IN A MODIFICATION OF  
CC THE 3'END OF THE TAIL FIBER GENE AND ALTERS THE HOST SPECIFICITY.  
CC -!- SIMILARITY: BELONGS TO THE SITE-SPECIFIC RECOMBINASE RESOLVASE  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way



CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

DR EMBL; X05926; CAA29365.1; -.  
DR InterPro; IPR001822; Recombinase.  
DR PROSITE; PS00397; RECOMBINASES\_1; PARTIAL.  
DR PROSITE; PS00398; RECOMBINASES\_2; PARTIAL.  
KW DNA recombination; DNA integration; DNA-binding; DNA invertase.  
FT NON\_TER 1 1  
FT DNA\_BIND <1 5 H-T-H MOTIF (PROBABLE).  
SQ SEQUENCE 22 AA; 2748 MW; 097E607032767C38 CRC64;

DNIV\_BPD10 Length: 22 December 22, 2002 19:30 Type: P Check: 9004 ..

1 YKKHPAKRTH IENDDRINQI DR

!!AA\_SEQUENCE 1.0  
ID FIBB\_ANTAM STANDARD; PRT; 21 AA.  
AC P14465;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Antilocapra americana (Pronghorn).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Antilocapridae; Antilocapra.  
OX NCBI\_TaxID=9891;  
RN [1]  
RP SEQUENCE.  
RA Gross G.A., Doolittle R.F.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
RL Arch. Biochem. Biophys. 122:674-684(1967).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2585 MW; FCE6183BE1F31627 CRC64;

FIBB\_ANTAM Length: 21 December 22, 2002 19:30 Type: P Check: 7016 ..

1 QPSYDYDEEE DDRAKLRD A R

!!AA\_SEQUENCE 1.0  
ID FIBB\_BISBO STANDARD; PRT; 21 AA.  
AC P14466;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Bison bonasus (European bison).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bison.  
OX NCBI\_TaxID=9902;  
RN [1]



RP SEQUENCE.  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2366 MW; 09EE75AF19E6363D CRC64;

FIBB\_BISBO Length: 21 December 22, 2002 19:30 Type: P Check: 7213 ..

1 EFPTDYDEGE DDRPKVGLGA R

!!AA\_SEQUENCE 1.0  
ID FIBB\_BUBBU STANDARD; PRT; 21 AA.  
AC P14467;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Bubalus bubalis (Domestic water buffalo).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bubalus.  
OX NCBI\_TaxID=89462;  
RN [1]  
RP SEQUENCE.  
RA Mross G.A., Doolittle R.F.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
RL Arch. Biochem. Biophys. 122:674-684(1967).  
RN [2]  
RP SEQUENCE.  
RC STRAIN=Italian breed;  
RX MEDLINE=76040091; PubMed=1180969;  
RA Balestrieri C., Colonna G., Irace G.;  
RT "Covalent structure of fibrinopeptides from buffaloes breeding in  
RT Italy.";  
RL Biochim. Biophys. Acta 405:517-521(1975).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2379 MW; 09EE75BE4729163D CRC64;

FIBB\_BUBBU Length: 21 December 22, 2002 19:30 Type: P Check: 7185 ..



1 QFPTDYDEGQ DDRPKLGLGA R

!!AA\_SEQUENCE 1.0  
ID FIBB\_CEREL STANDARD; PRT; 21 AA.  
AC P14468;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Cervus elaphus (Red deer), and  
OS Cervus elaphus nelsoni (American elk).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
OC Cervidae; Cervinae; Cervus.  
OX NCBI\_TaxID=9860, 9864;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=C.elaphus;  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=C.e.nelsoni;  
RA Mross G.A., Doolittle R.F.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
RL Arch. Biochem. Biophys. 122:674-684(1967).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2558 MW; FCEE745D98931627 CRC64;

FIBB\_CEREL Length: 21 December 22, 2002 19:30 Type: P Check: 6821 ..

1 QHSTDYDEEE EDRAKLHLDA R

!!AA\_SEQUENCE 1.0  
ID FIBB\_FELCA STANDARD; PRT; 20 AA.  
AC P14469;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE.  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS



CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma.  
FT PEPTIDE 1 20 FIBRINOPEPTIDE B.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2328 MW; A829E393B8F627D0 CRC64;

FIBB\_FELCA Length: 20 December 22, 2002 19:30 Type: P Check: 5816 ..

1 IIDYYDEGEE DRDVGVDAR

!!AA\_SEQUENCE 1.0  
ID FIBB\_LAMGL STANDARD; PRT; 19 AA.  
AC P14473;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Lama glama (Llama),  
OS Lama vicugna (Vicugna) (Vicugna vicugna), and  
OS Camelus dromedarius (Dromedary) (Arabian camel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
OX NCBI\_TaxID=9844, 9843, 9838;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=L.glama;  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=C.dromedarius;  
RX MEDLINE=67209145; PubMed=6033721;  
RA Doolittle R.F., Schubert D., Schwartz S.A.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I."  
RT Dromedary camel, mule deer, and cape buffalo.";  
RL Arch. Biochem. Biophys. 118:456-467(1967).  
RN [3]  
RP SEQUENCE.  
RC SPECIES=L.vicugna;  
RA Mross G.A., Doolittle R.F.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
RL Arch. Biochem. Biophys. 122:674-684(1967).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 19 FIBRINOPEPTIDE B.  
FT MOD\_RES 4 4 SULFATION.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2295 MW; E7EE6B6100568638 CRC64;

FIBB\_LAMGL Length: 19 December 22, 2002 19:30 Type: P Check: 4317 ..

1 ATDYDEEEDD RVKVRLDAR



!!AA\_SEQUENCE 1.0  
ID FIBB\_MUNMU STANDARD; PRT; 21 AA.  
AC P14475;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Muntiacus muntjak (Muntjak).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
OC Cervidae; Muntiacinae; Muntiacus.  
OX NCBI\_TaxID=9888;  
RN [1]  
RP SEQUENCE.  
RA Mross G.A., Doolittle R.F.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
RL Arch. Biochem. Biophys. 122:674-684(1967).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2514 MW; FCEE75188F0C1627 CRC64;

FIBB\_MUNMU Length: 21 December 22, 2002 19:30 Type: P Check: 6963 ..

1 QHSTDYDEVE DDRAKLHLDA R

!!AA\_SEQUENCE 1.0  
ID FIBB\_ODOHE STANDARD; PRT; 21 AA.  
AC P14476;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Odocoileus hemionus (Mule deer) (Black-tailed deer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
OC Cervidae; Odocoileinae; Odocoileus.  
OX NCBI\_TaxID=9872;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=67209145; PubMed=6033721;  
RA Doolittle R.F., Schubert D., Schwartz S.A.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.  
RT Dromedary camel, mule deer, and cape buffalo.";  
RL Arch. Biochem. Biophys. 118:456-467(1967).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.



KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2496 MW; FCF562C51A0C1627 CRC64;

FIBB\_ODOHE Length: 21 December 22, 2002 19:30 Type: P Check: 6856 ..

1 QHLADYDEVD DDRALKHLDA R

!!AA\_SEQUENCE 1.0  
ID FIBB\_RANTA STANDARD PRT; 21 AA.  
AC P14479;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Rangifer tarandus (Reindeer) (Caribou).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
OC Cervidae; Odocoileinae; Rangifer.  
OX NCBI\_TaxID=9870;  
RN [1]  
RP SEQUENCE.  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPRO002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 6 6 SULFATION.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2510 MW; FCF562C45F0C1627 CRC64;

FIBB\_RANTA Length: 21 December 22, 2002 19:30 Type: P Check: 6866 ..

1 QHLADYDEVE DDRALKHLDA R

!!AA\_SEQUENCE 1.0  
ID FIBB\_SYNCA STANDARD PRT; 21 AA.  
AC P14481;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Syncerus caffer (Cape buffalo).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Syncerus.  
OX NCBI\_TaxID=9970;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=67209145; PubMed=6033721;  
RA Doolittle R.F., Schubert D., Schwartz S.A.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.  
RT Dromedary camel, mule deer, and cape buffalo.";  
RL Arch. Biochem. Biophys. 118:456-467(1967).



CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2353 MW; 09EE75AF033B863D CRC64;

FIBB\_SYNCA Length: 21 December 22, 2002 19:30 Type: P Check: 7177 ..

1 QFPTDYDEGE DDRPKSGLGA R

!!AA\_SEQUENCE 1.0

ID GLU1\_ORENI STANDARD PRT; 36 AA.  
 AC P81026;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Glucagon I.  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;  
 OC Cichlidae; Oreochromis.  
 OX NCBI\_TaxID=8128;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95384941; PubMed=7656183;  
 RA Nguyen T.M., Wright J.R. Jr., Nielsen P.F., Conlon J.M.;  
 RT "Characterization of the pancreatic hormones from the Brockmann body  
 of the tilapia: implications for islet xenograft studies.";  
 RL Comp. Biochem. Physiol. 111C:33-44(1995).  
 CC -!- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES  
 CC THE BLOOD SUGAR LEVEL.  
 CC -!- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS  
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.  
 CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.  
 DR HSSP; P01274; 1GCN.  
 DR InterPro; IPR000532; Glucagon.  
 DR Pfam; PF00123; hormone2; 1.  
 DR PRINTS; PR00275; GLUCAGON.  
 DR SMART; SM00070; GLUCA; 1.  
 DR PROSITE; PS00260; GLUCAGON; 1.  
 KW Glucagon family; Hormone.  
 SQ SEQUENCE 36 AA; 4252 MW; 5AAFBF0254425756 CRC64;

GLU1\_ORENI Length: 36 December 22, 2002 19:30 Type: P Check: 382 ..

1 HSEGTFNSNDY SKYLEDRKAQ DFVRWLMNNK RSGAAE

!!AA\_SEQUENCE 1.0

ID LPGE\_ECOLI STANDARD PRT; 19 AA.  
 AC P33236;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Gef leader peptide.  
 GN GEFL OR B0018.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.



OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92048481; PubMed=1943701;  
RA Poulsen L.K., Refn A., Molin S., Andersson P.;  
RT "The gef gene from Escherichia coli is regulated at the level of  
RT translation.";  
RL Mol. Microbiol. 5:1639-1648(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; AE000112; AAC73129.1; ALT\_TERM.  
DR PIR; S16473; S16473.  
DR EcoGene; EG12074; gefL.  
KW Leader peptide; Complete proteome.  
SQ SEQUENCE 19 AA; 2259 MW; 19B3EDF371EB0BEB CRC64;

LPGE\_ECOLI Length: 19 December 22, 2002 19:30 Type: P Check: 4518 ..

1 MLNTCRVPLT DRKVKEKRA

!!AA\_SEQUENCE 1.0  
ID MY14\_EISFO STANDARD PRT; 14 AA.  
AC P46979;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Myoactive tetradecapeptide (ETP).  
OS Eisenia foetida (Common brandling worm) (Common dung-worm).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
OC Lumbricina; Lumbricidae; Eisenia.  
OX NCBI\_TaxID=6396;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Gut;  
RX MEDLINE=96087879; PubMed=8532604;  
RA Ukena K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,  
RA Nomoto K.;  
RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia  
foetida.";  
RL Peptides 16:995-999(1995).  
CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT  
CC MUSCLES.  
CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 14 14 AMIDATION.  
SQ SEQUENCE 14 AA; 1478 MW; CC9ABEF941CD91AD CRC64;

MY14\_EISFO Length: 14 December 22, 2002 19:30 Type: P Check: 7600 ..

1 GFKDGAADRI SHGF

!!AA\_SEQUENCE 1.0  
ID MY14\_PHEVI STANDARD PRT; 14 AA.  
AC P46980;



DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Myoactive tetradecapeptide (PTP).  
OS Pheretima vittata (Earthworm).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
OC Lumbricina; Megascolecidae; Pheretima.  
OX NCBI\_TaxID=46674;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Gut;  
RX MEDLINE=96087879; PubMed=8532604;  
RA Ukena K., Oumi T., Matsushima O., Ikeda T., Fujita T., Minakata H.,  
RA Nomoto K.;  
RT "A novel gut tetradecapeptide isolated from the earthworm, Eisenia  
RT foetida.";  
RL Peptides 16:995-999(1995).  
CC -!- FUNCTION: HAS A STIMULATIVE EFFECT ON THE CONTRACTION OF GUT  
CC MUSCLES.  
CC -!- SIMILARITY: TO INSECTS ALLATOTROPIN.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 14 14 AMIDATION.  
SQ SEQUENCE 14 AA; 1522 MW; DA40BEE67CCD91AD CRC64;

MY14\_PHEVI Length: 14 December 22, 2002 19:30 Type: P Check: 7729 ..

1 GFRDGSA DRI SHGF

!!AA\_SEQUENCE 1.0  
ID T2A\_PARTE STANDARD; PRT; 23 AA.  
AC Q27173;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Trichocyst matrix protein T2-A (Secretory granule protein T2-A)  
DE (TMP 2-A) (Fragment).  
GN T2A.  
OS Paramecium tetraurelia.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;  
OC Paramecium.  
OX NCBI\_TaxID=5888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D4-2;  
RX MEDLINE=96059477; PubMed=7579685;  
RA Madeddu L., Gautier M.-C., Vayssie L., Houari A., Sperling L.;  
RT "A large multigene family codes for the polypeptides of the  
RT crystalline trichocyst matrix in Paramecium.";  
RL Mol. Biol. Cell 6:649-659(1995).  
RN [2]  
RP PARTIAL SEQUENCE.  
RC STRAIN=D4-2;  
RX MEDLINE=95119139; PubMed=7819344;  
RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;  
RT "Protein processing and morphogenesis of secretory granules in  
RT Paramecium.";  
RL Biochimie 76:329-335(1994).  
CC -!- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE  
CC TRICHOCYST MATRIX.  
CC -!- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY  
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,  
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.  
CC -!- SIMILARITY: BELONGS TO THE TMP FAMILY.  
CC -!- DATABASE: NAME=Protein Spotlight;  
CC NOTE=Issue 3 of October 2000;  
CC WWW="<http://www.expasy.org/spotlight/articles/sptlt003.html>".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way.



CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; U27509; AAA92609.1; -.

KW Polyprotein; Structural protein; Multigene family.

FT NON\_TER 1 1

FT NON\_TER 23 23

SQ SEQUENCE 23 AA; 2720 MW; 08972B1F2759BDB7 CRC64;

T2A\_PARTE Length: 23 December 22, 2002 19:30 Type: P Check: 946 ..

1 DPLDRLLSTL TDLEDTRYVAE QKE

!!AA\_SEQUENCE 1.0

ID T2B\_PARTE STANDARD; PRT; 23 AA.

AC Q27174;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Trichocyst matrix protein T2-B (Secretory granule protein T2-B)  
(TMP 2-B) (Fragment).

GN T2B.

OS Paramecium tetraurelia.

OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;

OC Paramecium.

OX NCBI\_TaxID=5888;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D4-2;

RX MEDLINE=96059477; PubMed=7579685;

RA Madeddu L., Gautier M.-C., Vayssie L., Houari A., Sperling L.;

RT "A large multigene family codes for the polypeptides of the  
crystalline trichocyst matrix in Paramecium.";

RL Mol. Biol. Cell 6:649-659(1995).

RN [2]

RP PARTIAL SEQUENCE.

RC STRAIN=D4-2;

RX MEDLINE=95119139; PubMed=7819344;

RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;

RT "Protein processing and morphogenesis of secretory granules in  
Paramecium.";

RL Biochimie 76:329-335(1994).

CC --!- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE  
CC TRICHOCYST MATRIX.

CC --!- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY  
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,  
CC READY TO RAPIDLY RESPOND TO AN EXOCYTOTIC STIMULUS.

CC --!- SIMILARITY: BELONGS TO THE TMP FAMILY.

CC --!- DATABASE: NAME=Protein Spotlight;

CC NOTE=Issue 3 of October 2000;

CC WWW="<http://www.expasy.org/spotlight/articles/sptlt003.html>".

CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; U27510; AAA92610.1; -.

KW Polyprotein; Structural protein; Multigene family.

FT NON\_TER 1 1

FT NON\_TER 23 23

SQ SEQUENCE 23 AA; 2706 MW; 184D4B1F2759BDB7 CRC64;

T2B\_PARTE Length: 23 December 22, 2002 19:30 Type: P Check: 1016 ..

1 DPLDRLVSTL TDLEDTRYVAE QKE



!!AA\_SEQUENCE 1.0  
ID Q9ZG33 PRELIMINARY; PRT; 17 AA.  
AC Q9ZG33;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE Subtilisin/chymotrypsin inhibitor (Fragment).  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L2 434B;  
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
RT "Gene identification of Chlamydia trachomatis by random DNA  
RT sequencing.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF087336; AAD04110.1; -.  
FT NON\_TERMINAL 1 1  
FT NON\_TERMINAL 17 17  
SQ SEQUENCE 17 AA; 2043 MW; 4FEE704EE041E120 CRC64;

Q9ZG33 Length: 17 December 22, 2002 19:30 Type: P Check: 1658 ..

1 NFLQPFCR RKNSFRI

!!AA\_SEQUENCE 1.0  
ID Q9KIL6 PRELIMINARY; PRT; 24 AA.  
AC Q9KIL6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE F81r (Fragment).  
GN F81R.  
OS Streptomyces coelicolor A3(2).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=100226;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M145;  
RA Kormanec J., Sevcikova B., Homerova D.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF230491; AAF82064.1; -.  
FT NON\_TERMINAL 24 24  
SQ SEQUENCE 24 AA; 2524 MW; FDA1B812ED67583B CRC64;

Q9KIL6 Length: 24 December 22, 2002 19:30 Type: P Check: 2042 ..

1 MAADRTGTDE ATAERALGSR APEF

!!AA\_SEQUENCE 1.0  
ID Q9X9U4 PRELIMINARY; PRT; 25 AA.  
AC Q9X9U4;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE ADP glucose pyrophosphorylase (Fragment).  
GN GLGC.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Bruton C.J.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);



RX MEDLINE=98062210; PubMed=9401038;  
RA Martin M., Schneider D., Bruton C.J., Chater K.F., Hardisson C.;  
RT "A glgC Gene Essential Only for the First of Two Spatially Distinct  
Phases of Glycogen Synthesis in Streptomyces coelicolor A3(2).";  
RL J. Bacteriol. 179:7784-7789(1997).  
DR EMBL; AJ243803; CAB50742.1; -.  
DR InterPro; IPR001825; NTP\_transferase.  
DR Pfam; PF00483; NTP\_transferase; 1.  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 2566 MW; 84B94A728A41D25C CRC64;

Q9X9U4 Length: 25 December 22, 2002 19:30 Type: P Check: 4039 ..

1 MLGIVLAGGE GKRLMPLTAD RAKPA

!!AA\_SEQUENCE 1.0  
ID Q49137 PRELIMINARY; PRT; 32 AA.  
AC Q49137;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE MxaS protein (Fragment).  
GN MXAS.  
OS Methylobacterium extorquens.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Methylobacterium group; Methylobacterium.  
OX NCBI\_TaxID=408;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AM1;  
RX MEDLINE=96074311; PubMed=7592474;  
RA Morris C.J., Kim Y.M., Perkins K.E., Lidstrom M.E.;  
RT "Identification and nucleotide sequences of mxaA, mxaC, mxaK, mxaL,  
and mxA genes from Methylobacterium extorquens AM1.";  
RL J. Bacteriol. 177:6825-6831(1995).  
DR EMBL; L41608; AAA85567.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 32 AA; 3700 MW; EB03BB62B3C56ED5 CRC64;

Q49137 Length: 32 December 22, 2002 19:30 Type: P Check: 9647 ..

1 ALRRICAPFA RPPFRLADRF DAEALSRHLM TT

!!AA\_SEQUENCE 1.0  
ID Q53914 PRELIMINARY; PRT; 9 AA.  
AC Q53914;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE CurD protein (Fragment).  
OS Streptomyces cyaneus (Streptomyces curacoi).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1904;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bergh S.T., Uhlen M.;  
RT "Cloning, analysis and heterologous expression of the polyketides  
synthesis genes of Streptomyces curacoi.";  
RL Submitted (JUN-1990) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; M33704; AAA26724.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 1027 MW; 995BDDDC4140AB1 CRC64;

Q53914 Length: 9 December 22, 2002 19:30 Type: P Check: 3396 ..

1 ITDRRAAQP

!!AA\_SEQUENCE 1.0  
ID Q50476 PRELIMINARY; PRT; 13 AA.  
AC Q50476;



DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Catalase.  
GN KATG.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L10373;  
RA Cockerill F.R., Uhl J.R., Temesgen Z., Zhang Y., Stockman L.,  
RA Roberts G.D., Williams D.L., Kline B.C.;  
RT "Rapid Identification of a point mutation of the Mycobacterium  
RT tuberculosis catalase-peroxidase (katG) gene associated with isoniazid  
RT resistance.";  
RL Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U06263; AAB59976.1; -.  
SQ SEQUENCE 13 AA; 1564 MW; 2F39A45EFE994777 CRC64;

Q50476 Length: 13 December 22, 2002 19:30 Type: P Check: 7152 ..

1 MPEQHPPITD RSR

!!AA\_SEQUENCE 1.0  
ID 005422 PRELIMINARY; PRT; 38 AA.  
AC 005422;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Phytoene dehydrogenase (Fragment).  
GN CRTI.  
OS Mycobacterium marinum.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1781;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M;  
RX MEDLINE=97440138; PubMed=9294446;  
RA Ramakrishnan L., Tran H.T., Federspiel N.A., Falkow S.;  
RT "A crtB homolog essential for photochromogenicity in Mycobacterium  
RT marinum: isolation, characterization, and gene disruption via  
RT homologous recombination.";  
RL J. Bacteriol. 179:5862-5868(1997).  
DR EMBL; U92075; AAB71427.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 38 AA; 3986 MW; 6E46332707CCDCAB CRC64;

005422 Length: 38 December 22, 2002 19:30 Type: P Check: 6757 ..

1 VPGVGVPPTL ISGRLAADRI TGNTTRSIRH LDLKAQLS

!!AA\_SEQUENCE 1.0  
ID Q93A12 PRELIMINARY; PRT; 22 AA.  
AC Q93A12;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Putative ribitol/glucose oxidoreductase (Fragment).  
GN SDRA2.  
OS Thiobacillus ferrooxidans.  
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
OX NCBI\_TaxID=920;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC33020;  
RA Bruscella P., Levican G., Ratouchniak J., Holmes D., Bonnefoy V.;  
RT "A second operon encoding a bcl complex in Acidithiobacillus  
RT ferrooxidans.";



RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ311888; CAC44744.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 22 AA; 2464 MW; 3725F8E43BBA75C2 CRC64;

Q93A12 Length: 22 December 22, 2002 19:30 Type: P Check: 9419 ..

1 AQNFINPDLD DRTEKDLATS TV

!!AA\_SEQUENCE 1.0  
ID Q93A08 PRELIMINARY; PRT; 12 AA.  
AC Q93A08;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE ResB protein (Fragment).  
GN RESB.  
OS Thiobacillus ferrooxidans.  
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.  
OX NCBI\_TaxID=920;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC33020;  
RA Levican G., Bruscella P., Guacunano M., Inostroza C., Jedlicki E.,  
RA Bonnefoy V., Holmes D.S.;  
RT "Characterization of the pet and res operons of Acidithiobacillus  
RT ferrooxidans.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ413194; CAC88360.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 12 AA; 1405 MW; 886AB7DF1E13240A CRC64;

Q93A08 Length: 12 December 22, 2002 19:30 Type: P Check: 5753 ..

1 QSQDDRKENN DG

!!AA\_SEQUENCE 1.0  
ID Q03977 PRELIMINARY; PRT; 19 AA.  
AC Q03977;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Alpha-acceptor polypeptide M15 (Fragment).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DH5 ALPHA;  
RX MEDLINE=93083990; PubMed=1339377;  
RA Prentki P.C.;  
RT "Nucleotide sequence of the classical lacZ deletion delta M15.";  
RL Gene 122:231-232(1992).  
DR EMBL; X58252; CAA41206.1; -.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2092 MW; 8324315E003AA053 CRC64;

Q03977 Length: 19 December 22, 2002 19:30 Type: P Check: 4923 ..

1 MTMITDSLAV VARTDRPSQ

!!AA\_SEQUENCE 1.0  
ID Q99094 PRELIMINARY; PRT; 36 AA.  
AC Q99094;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last annotation update)  
DE MKAA protein (Fragment).  
OS Salmonella typhimurium.  
OG Plasmid pYA426.



OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gulig P.A., Chioldo V.A.;  
RL Infect. Immun. 58:2651-2658(1991).  
DR EMBL; M64295; AAA27272.1; -.  
KW Plasmid.  
FT NON TER 1 1  
SQ SEQUENCE 36 AA; 4121 MW; C3D43D1C622EBAB7 CRC64;

Q99094 Length: 36 December 22, 2002 19:30 Type: P Check: 1642 ..

1 GSQDFASQLS KLRLSDDRTA DTNRIKRIIN MRVLNS

!!AA\_SEQUENCE 1.0  
ID Q9NZ25 PRELIMINARY; PRT; 20 AA.  
AC Q9NZ25;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE ATP7B (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wu Z.Y., Wang N., MuRong S.X., Lin M.T., Fang L.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF220215; AAF67661.1; -.  
DR InterPro; IPR001757; ATPase\_E1-E2.  
DR Pfam; PF00122; E1-E2\_ATPase; 1.  
FT NON TER 1 1  
FT VARIANT 19 19 V -> I.  
FT NON TER 20 20  
SQ SEQUENCE 20 AA; 2282 MW; 00268E2C1E0692E3 CRC64;

Q9NZ25 Length: 20 December 22, 2002 19:30 Type: P Check: 6136 ..

1 APIQLADRF SGYFVPFIVI

!!AA\_SEQUENCE 1.0  
ID Q15244 PRELIMINARY; PRT; 40 AA.  
AC Q15244;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE Phosphoribosylpyrophosphate synthetase isoform I (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BLOOD;  
RA Ishizuka T.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BLOOD;  
RX MEDLINE=92223087; PubMed=1314091;  
RA Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shimada H.,  
RA Nagatake N., Tatibana M.;  
RT "Promoter regions of the human X-linked housekeeping genes PRPS1 and  
RT PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II  
RT isoforms.";  
RL Biochim. Biophys. Acta 1130:139-148(1992).  
DR EMBL; D28133; BAA05675.1; -.  
DR HSSP; P14193; 1DKU.



FT NON\_TER 40 40  
SQ SEQUENCE 40 AA; 4446 MW; 57C2B4011860B098 CRC64;  
Q15244 Length: 40 December 22, 2002 19:30 Type: P Check: 2868 ..  
1 MPNIKIFSGS SHQDLSQKIA DRLGLELGKV VTKKFSNQET

!!AA\_SEQUENCE 1.0  
ID Q15245 PRELIMINARY; PRT; 40 AA.  
AC Q15245;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE Phosphoribosylpyrophosphate synthetase isoform II (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BLOOD;  
RA Ishizuka T.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BLOOD;  
RX MEDLINE=92223087; PubMed=1314091;  
RA Ishizuka T., Iizasa T., Taira M., Ishijima S., Sonoda T., Shimada H.,  
RA Nagatake N., Tatibana M.;  
RT "Promoter regions of the human X-linked housekeeping genes PRPS1 and  
RT PRPS2 encoding phosphoribosylpyrophosphate synthetase subunit I and II  
RT isoforms.";  
RL Biochim. Biophys. Acta 1130:139-148(1992).  
DR EMBL; D28134; BAA05676.1; -.  
FT NON\_TER 40 40  
SQ SEQUENCE 40 AA; 4431 MW; CEC2B31686083EFF CRC64;

Q15245 Length: 40 December 22, 2002 19:30 Type: P Check: 3314 ..

1 MPNIVLFSGS SHQDLSQRVA DRLGLELGKV VTKKFSNQET

!!AA\_SEQUENCE 1.0  
ID Q9BQV8 PRELIMINARY; PRT; 34 AA.  
AC Q9BQV8;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE L-isoaspartyl/D-aspartyl methyltransferase (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94354847; PubMed=8074695;  
RA Tsai W., Clarke S.;  
RT "Amino acid polymorphisms of the human L-isoaspartyl/D-aspartyl  
RT methyltransferase involved in protein repair.";  
RL Biochem. Biophys. Res. Commun. 203:491-497(1994).  
DR EMBL; S73902; AAC60639.2; -.  
DR InterPro; IPR000682; Pcm.  
DR Pfam; PF01135; PCMT; 1.  
KW Methyltransferase; Transferase.  
FT NON\_TER 1 1  
FT NON\_TER 34 34  
SQ SEQUENCE 34 AA; 3873 MW; EEB18E55F49BF377 CRC64;

Q9BQV8 Length: 34 December 22, 2002 19:30 Type: P Check: 5960 ..

1 NGIIKTDKVF EVMLATDRSH YAKCNPYMD5 PQSI



!!AA\_SEQUENCE 1.0  
ID Q9UCI1 PRELIMINARY; PRT; 15 AA.  
AC Q9UCI1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Tropomyosin (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93195352; PubMed=8450225;  
RA Das K.M., Dasgupta A., Mandal A., Geng X.;  
RT "Autoimmunity to cytoskeletal protein tropomyosin. A clue to the  
RT pathogenetic mechanism for ulcerative colitis.";  
RL J. Immunol. 150:2487-2493(1993).  
DR InterPro; IPR000533; Tropomyosin.  
DR Pfam; PF00261; Tropomyosin; 1.  
SQ SEQUENCE 15 AA; 1802 MW; 7A0993CA5A54254C CRC64;

Q9UCI1 Length: 15 December 22, 2002 19:30 Type: P Check: 8961 ..

1 HIAEDADRKY EEVAR

!!AA\_SEQUENCE 1.0  
ID Q96F68 PRELIMINARY; PRT; 28 AA.  
AC Q96F68;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Unknown (Protein for IMAGE:4563468) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; BC011565; AAH11565.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 28 AA; 2841 MW; 5799D138245D3951 CRC64;

Q96F68 Length: 28 December 22, 2002 19:30 Type: P Check: 588 ..

1 VSQPGSCRHG ADRVGHVGQR AGAGVRPE

!!AA\_SEQUENCE 1.0  
ID Q9NBB1 PRELIMINARY; PRT; 26 AA.  
AC Q9NBB1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Stretchin-MLCK (Fragment).  
GN STRN-MLCK OR CG8304 OR CG18255.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephdroioidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Champagne M.B., Edwards K.A., Erickson H.P., Kiehart D.P.;  
RT "Drosophila stretchin-MLCK is a Novel Member of the Titin/Myosin Light  
RT Chain Kinase Family.";  
RL J. Mol. Biol. 0:0-0(2000).  
DR EMBL; AF257309; AAF90127.1; -.  
DR FlyBase; FBgn0013988; Strn-Mlck.



FT NON\_TER 1 1  
SQ SEQUENCE 26 AA; 3004 MW; 2F2CB3A55E7FF033 CRC64;  
Q9NBB1 Length: 26 December 22, 2002 19:30 Type: P Check: 6662 ..  
1 VVARNNFGTD RIFVTVTIKI PKKKEE  
!!AA\_SEQUENCE 1.0  
ID Q9GU45 PRELIMINARY; PRT; 27 AA.  
AC Q9GU45;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Homeobox protein Sycox2 (Fragment).  
GN SYCOX2.  
OS Sycon raphanus.  
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;  
OC Sycettidae.  
OX NCBI\_TaxID=56443;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20476455; PubMed=11020308;  
RA Manuel M., Le Parco Y.;  
RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon raphanus.";  
RL Mol. Phylogenet. Evol. 17:97-107(2000).  
DR EMBL; AF197140; AAG28510.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2986 MW; CA58FA11B629E3B7 CRC64;  
Q9GU45 Length: 27 December 22, 2002 19:30 Type: P Check: 9325 ..  
1 KTSKYLAPTT RQALAARLGL TDRQVKV  
!!AA\_SEQUENCE 1.0  
ID Q9GU44 PRELIMINARY; PRT; 27 AA.  
AC Q9GU44;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Homeobox protein Sycox3 (Fragment).  
GN SYCOX3.  
OS Sycon raphanus.  
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;  
OC Sycettidae.  
OX NCBI\_TaxID=56443;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20476455; PubMed=11020308;  
RA Manuel M., Le Parco Y.;  
RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon raphanus.";  
RL Mol. Phylogenet. Evol. 17:97-107(2000).  
DR EMBL; AF197141; AAG28511.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3014 MW; CA58E2803F29E3B7 CRC64;  
Q9GU44 Length: 27 December 22, 2002 19:30 Type: P Check: 9332 ..  
1 RTSKYLAPTT RQALAARLGL TDRQVKV  
!!AA\_SEQUENCE 1.0  
ID Q9XZZ6 PRELIMINARY; PRT; 23 AA.  
AC Q9XZZ6;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)



DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE Calmodulin (Fragment).  
GN CAM.  
OS Littorina saxatilis.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Mesogastropoda; Littorinoidea; Littorinidae; Littorina.  
OX NCBI\_TaxID=31220;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wilding C.S., Grahame J., Mill P.J.;  
RT "Molecular characterisation of calmodulin intron variation in  
RT Littorina (Gastropoda: Prosobranchia) species.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ133335; CAB44224.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2815 MW; 5762B9C4E9886A06 CRC64;

Q9XZZ6 Length: 23 December 22, 2002 19:30 Type: P Check: 1592 ..

1 LQDMINEVDA DRQRDDRLPR VPY

!!AA\_SEQUENCE 1.0  
ID Q94681 PRELIMINARY; PRT; 25 AA.  
AC Q94681;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PmHbox1 (Fragment).  
OS Polyandrocarpa misakiensis.  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
OC Stolidobranchia; Styelidae; Polyandrocarpa.  
OX NCBI\_TaxID=7723;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USHIMADO;  
RA Fujiwara S., Kawamura K.;  
RT "Cloning of homeobox-containing genes from the budding ascidian  
RT Polyandrocarpa misakiensis.";  
RL Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; D88274; BAA13570.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 3074 MW; D7CBE263F4B30A51 CRC64;

Q94681 Length: 25 December 22, 2002 19:30 Type: P Check: 5282 ..

1 HFNQYLCRER RQEVAKAVNL TDRQV

!!AA\_SEQUENCE 1.0  
ID Q23749 PRELIMINARY; PRT; 27 AA.  
AC Q23749;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Ovx1 ortholog homeobox (Fragment).  
GN CTS-OVX1.  
OS Ctenodrilus serratus.  
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;  
OC Terebellida; Ctenodrilidae; Ctenodrilus.  
OX NCBI\_TaxID=40316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94356262; PubMed=7915607;  
RA Dick M.H., Buss L.W.;  
RT "A PCR-based survey of homeobox genes in Ctenodrilus serratus  
RT (Annelida: Polychaeta).";  
RL Mol. Phylogenetic Evol. 3:146-158(1994).



RN [2]  
RP SEQUENCE FROM N.A.  
RA Dick M.H., Buss L.W.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U26634; AAC46856.1; -.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2879 MW; 4CCE222303AA0102 CRC64;

Q23749 Length: 27 December 22, 2002 19:30 Type: P Check: 8475 ..

1 PQQEILVTDG TIAHRASPET DRGSGEN

!!AA\_SEQUENCE 1.0  
ID Q25482 PRELIMINARY; PRT; 27 AA.  
AC Q25482;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Anthox1.Ms (Fragment).  
GN ANTHOX1.MS.  
OS Metridium senile (Brown sea anemone) (Frilled sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
OC Nynantheae; Metridiidae; Metridium.  
OX NCBI\_TaxID=6116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97435515; PubMed=9290214;  
RA Finnerty J.R., Martindale M.Q.;  
RT "Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of  
RT Nematostella vectensis and Metridium senile.";  
RL Biol. Bull. 193:62-76(1997).  
DR EMBL; U42727; AAA86626.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEOBOX.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3314 MW; E6B0D191EE3905CF CRC64;

Q25482 Length: 27 December 22, 2002 19:30 Type: P Check: 9269 ..

1 HFNHFLTKER RSEMATQLNL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID Q94458 PRELIMINARY; PRT; 27 AA.  
AC Q94458;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE CHv-Hb11 protein (Fragment).  
GN CHV-HB11.  
OS Chaetopterus variopedatus.  
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;  
OC Spionida; Chaetopteridae; Chaetopterus.  
OX NCBI\_TaxID=34590;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Irvine S.M., Warinner S.A., Hunter J.D., Martindale M.Q.;  
RT "A survey of homeobox genes in Chaetopterus variopedatus and analysis  
RT of polychaete homeodomains.";  
RL Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U68283; AAB16992.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3351 MW; 74E5C77ED08CDEE4 CRC64;

Q94458 Length: 27 December 22, 2002 19:30 Type: P Check: 8578 ..

1 NQKKFIEKKD RDRISNEIGL DDRQIKY



!!AA\_SEQUENCE 1.0  
ID 096892 PRELIMINARY; PRT; 27 AA.  
AC 096892;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Caudal (Fragment).  
GN CAD.  
OS Sacculina carcinia.  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;  
OC Rhizocephala; Kentrogonida; Sacculinidae; Sacculina.  
OX NCBI\_TaxID=51650;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98334561; PubMed=9667986;  
RA Mouchel-Vielh E., Rigolot C., Gibert J.M., Deutsch J.S.;  
RT "Molecules and the body plan: the Hox genes of Cirripedes  
(Crustacea).";  
RL Mol. Phylogenetic Evol. 9:382-389(1998).  
DR EMBL; U79471; AAD00342.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEOBOX.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3289 MW; 1CDC2E4CCA04E00F CRC64;

096892 Length: 27 December 22, 2002 19:30 Type: P Check: 9212 ..

1 RFNNYITIKR KLELSRILGL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID P91901 PRELIMINARY; PRT; 27 AA.  
AC P91901;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Homeobox protein BHox26 (Fragment).  
OS Beroe ovata.  
OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroida; Beroidae; Beroe.  
OX NCBI\_TaxID=10201;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Witchell H.J., Meech R.W.;  
RT "BHox26 - Beroe Homeobox.";  
RL Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U89381; AAB49471.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEOBOX.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

P91901 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1 LFNMYLTRER RLEISKSINL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID 002602 PRELIMINARY; PRT; 27 AA.  
AC 002602; O02535;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Homeobox protein BHox35 (Fragment).  
OS Beroe ovata.  
OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroida; Beroidae; Beroe.  
OX NCBI\_TaxID=10201;  
RN [1]  
RP SEQUENCE FROM N.A.



RA Witchel H.J., Meech R.W.;  
RT "BHox35 - Beroe Homeobox.";  
RL Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U89383; AAB49473.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

002602 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..

1 LFNMYLTRER RLEISRGVNL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID O18615 PRELIMINARY; PRT; 34 AA.  
AC O18615;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE FTZ-F1 (Fragment).  
OS Artemia salina (Brine shrimp).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;  
OC Artemiidae; Artemia.  
OX NCBI\_TaxID=85549;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97338099; PubMed=9192646;  
RA Escriva H., Safi R., Hanni C., Langlois M.C., Saumitou-Laprade P.,  
RA Stehelin D., Capron A., Pierce R., Laudet V.;  
RT "Ligand binding was acquired during evolution of nuclear receptors.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:6803-6808(1997).  
CC -!!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -!!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
DR EMBL; U93418; AAC83395.1; -.  
DR InterPro; IPR001628; Znf\_C4steroid.  
DR Pfam; PF00105; zf-C4; 1.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00399; ZnF\_C4; 1.  
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
KW Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 34 34  
SQ SEQUENCE 34 AA; 4132 MW; 1AE9E6A0D63D4AEF CRC64;

018615 Length: 34 December 22, 2002 19:30 Type: P Check: 5276 ..

1 FKRTVQNKKV YTGMADRSCH IDKSQRKRCP PCRF

!!AA\_SEQUENCE 1.0  
ID Q9V973 PRELIMINARY; PRT; 17 AA.  
AC Q9V973;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE CG12485 protein.  
GN CG12485.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydriidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,



RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin R.K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003791; AAF57423.1; -.  
DR FlyBase; FBgn0034547; CG12485.  
SQ SEQUENCE 17 AA; 1907 MW; 47D598D29F4860DE CRC64;

Q9V973 Length: 17 December 22, 2002 19:30 Type: P Check: 1645 ..

1 MELSSTDGRP HHAQQGR

!!AA\_SEQUENCE 1.0  
ID Q9TWR1 PRELIMINARY; PRT; 21 AA.  
AC Q9TWR1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE Cysteine proteinase (Fragment).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Bombycoidea; Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94086476; PubMed=8262908;  
RA Takahashi S.Y., Yamamoto Y., Shionoya Y., Kageyama T.;  
RT "Cysteine proteinase from the eggs of the silkworm, *Bombyx mori*:  
RT identification of a latent enzyme and characterization of activation  
RT and proteolytic processing in vivo and in vitro.";  
RL J. Biochem. 114:267-272(1993).  
SQ SEQUENCE 21 AA; 2379 MW; 9232784B6D6AE4B1 CRC64;

Q9TWR1 Length: 21 December 22, 2002 19:30 Type: P Check: 8106 ..

1 XPEQVDDRKH GAVTDDKXXQ X

!!AA\_SEQUENCE 1.0  
ID O62575 PRELIMINARY; PRT; 26 AA.



AC 062575;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hox protein MYX4 (Fragment).  
GN HOX OR MYX4.  
OS Tetracapsula bryozoides.  
OC Eukaryota; Metazoa; Myxozoa; Myxozoa incertae sedis; Tetracapsula.  
OX NCBI\_TaxID=75694;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98196665; PubMed=9537319;  
RA Anderson C.L., Canning E.U., Okamura B.;  
RT "A triploblast origin for Myxozoa?";  
RL Nature 392:346-347(1998).  
DR EMBL; AJ005124; CAA06387.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR ProDom; PD000010; Homeobox; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 3212 MW; E38651AD1657BD2C CRC64;

062575 Length: 26 December 22, 2002 19:30 Type: P Check: 7242 ..

1 LFNMYLSRER RLEISKSIDL TDRQVK

!!AA\_SEQUENCE 1.0  
ID Q61676 PRELIMINARY; PRT; 27 AA.  
AC Q61676;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE HOX11-D125 protein (D125) (Fragment).  
GN HOX11-D125 OR D125.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydriidae; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93281593; PubMed=8099440;  
RA Dear T.N., Sanchez-Garcia I., Rabbitts T.H.;  
RT "The HOX11 gene encodes a DNA-binding nuclear transcription factor  
belonging to a distinct family of homeobox genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).  
DR EMBL; L08618; AAA28612.1; -. ✓  
DR FlyBase; FBgn0010394; Hox11-D125.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2937 MW; CABBC2F876F2B3A3 CRC64;

Q61676 Length: 27 December 22, 2002 19:30 Type: P Check: 8698 ..

1 LYQKYLSPAD RDEIAASLGL SNAQVIT

!!AA\_SEQUENCE 1.0  
ID Q9MZW8 PRELIMINARY; PRT; 29 AA.  
AC Q9MZW8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE Von Willebrand factor (Fragment).  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.



OX NCBI\_TaxID=9600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20072937; PubMed=10603266;  
RA Chaves R., Sampaio I., Schneider M.P., Schneider H., Page S.L.,  
RA Goodman M.;  
RT "The place of Callimico goeldii in the Callitrichine phylogenetic  
RT tree: evidence from von Willebrand factor gene intron II sequences.";  
RL Mol. Phylogenet. Evol. 13:392-404(1999).  
DR EMBL; AF092833; AAF77601.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 29 29  
SQ SEQUENCE 29 AA; 3227 MW; 60F1190C8F227D03 CRC64;

Q9MZ8 Length: 29 December 22, 2002 19:30 Type: P Check: 2931 ..

1 GRDCQDHFSF IVIETVQCAD DRDAVCTRS

!!AA\_SEQUENCE 1.0  
ID Q29394 PRELIMINARY; PRT; 23 AA.  
AC Q29394;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Growth hormone receptor (Fragment).  
GN GHR.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97049323; PubMed=8894053;  
RA Venta P.J., Brouillette J.A., Yuzbasiyan-Gurkan V., Brewer G.J.;  
RT "Gene-specific universal mammalian sequence-tagged sites: application  
RT to the canine genome.",  
RL Biochem. Genet. 34:321-341(1996).  
DR EMBL; L77673; AAA97423.1; -.  
KW Receptor.  
FT NON\_TER 1 1  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2620 MW; 85442F0EC46D5D08 CRC64;

Q29394 Length: 23 December 22, 2002 19:30 Type: P Check: 807 ..

1 DDLDEKTEGS DTDRLLSNDH EKS

!!AA\_SEQUENCE 1.0  
ID Q9TRS7 PRELIMINARY; PRT; 24 AA.  
AC Q9TRS7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Vitronectin (Fragment).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92207982; PubMed=1372829;  
RA Nakashima N., Miyazaki K., Ishikawa M., Yatohgo T., Ogawa H.,  
RA Uchibori H., Matsumoto I., Seno N., Hayashi M.;  
RT "Vitronectin diversity in evolution but uniformity in ligand binding  
RT and size of the core polypeptide.",  
RL Biochim. Biophys. Acta 1120:1-10(1992).  
DR InterPro; IPR000886; ER\_target.  
DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
FT NON\_TER 1 1  
FT NON\_TER 24 24  
SQ SEQUENCE 24 AA; 2745 MW; 94F0054986FE1215 CRC64;



Q9TRS7 Length: 24 December 22, 2002 19:30 Type: P Check: 2610 ..

1 AQESXKGRVT EGFNADRQQ QDEL

!!AA\_SEQUENCE 1.0  
ID Q9TQQ9 PRELIMINARY; PRT; 15 AA.  
AC Q9TQQ9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE Glutamate dehydrogenase isoform I (EC 1.4.1.2) (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96061967; PubMed=7588764;  
RA Cho S.W., Lee J., Choi S.Y.;  
RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine  
RT brain.";  
RL Eur. J. Biochem. 233:340-346(1995).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE=96043916; PubMed=7581004;  
RA Lee J., Kim S.W., Cho S.W.;  
RT "A novel glutamate dehydrogenase from bovine brain: purification and  
RT characterization.";  
RL Biochem. Mol. Biol. Int. 36:1087-1096(1995).  
SQ SEQUENCE 15 AA; 1754 MW; 65F7CD91023AEEBA CRC64;

Q9TQQ9 Length: 15 December 22, 2002 19:30 Type: P Check: 8644 ..

1 EEAAADREDD PNFFK

!!AA\_SEQUENCE 1.0  
ID Q9TR40 PRELIMINARY; PRT; 15 AA.  
AC Q9TR40;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE Glutamate dehydrogenase isoform II (EC 1.4.1.2) (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96061967; PubMed=7588764;  
RA Cho S.W., Lee J., Choi S.Y.;  
RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine  
RT brain.";  
RL Eur. J. Biochem. 233:340-346(1995).  
SQ SEQUENCE 15 AA; 1724 MW; 65F7CD91023AF925 CRC64;

Q9TR40 Length: 15 December 22, 2002 19:30 Type: P Check: 8661 ..

1 VEAAADREDD PNFFK

!!AA\_SEQUENCE 1.0  
ID Q9TRF2 PRELIMINARY; PRT; 33 AA.  
AC Q9TRF2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE VAMP/synaptobrevin-2 (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;



OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93374072; PubMed=8365494;  
RA Horikawa H.P., Saisu H., Ishizuka T., Sekine Y., Tsugita A., Odani S.,  
RA Abe T.;  
RL FEBS Lett. 330:236-240(1993).  
DR InterPro; IPR001388; Synaptobrevin.  
DR Pfam; PF00957; synaptobrevin; 1.  
DR ProDom; PD001229; Synaptobrevin; 1.  
DR PROSITE; PS00417; SYNAPTOBREVIN; 1.  
SQ SEQUENCE 33 AA; 3672 MW; 584386688353EC55 CRC64;

Q9TRF2 Length: 33 December 22, 2002 19:30 Type: P Check: 1454 ..

1 XVNVDKVLER DQKLSELDDR ADALQAGASQ FET

!!AA\_SEQUENCE 1.0  
ID Q37112 PRELIMINARY; PRT; 22 AA.  
AC Q37112;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE ORF22.  
OS Pinus thunbergii (Green pine) (Japanese black pine).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_TaxID=3350;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92212283; PubMed=1557027;  
RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,  
RA Wakasugi T., Sugiura M.;  
RT "Chloroplast DNA of black pine retains a residual inverted repeat  
RT lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and  
RT trnH and the absence of rps16.";  
RL Mol. Genet. 232:206-214(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95094312; PubMed=8001170;  
RA Tsudzuki J., Ito S., Tsudzuki T., Wakasugi T., Sugiura M.;  
RT "A new gene encoding tRNA(Pro) (GGG) is present in the chloroplast  
RT genome of black pine: a compilation of 32 tRNA genes from black pine  
RT chloroplasts.";  
RL Curr. Genet. 26:153-158(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95024047; PubMed=7937893;  
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,  
RA Sugiura M.;  
RT "Loss of all ndh genes as determined by sequencing the entire  
RT chloroplast genome of the black pine *Pinus thunbergii*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Sugiura M.;  
RL Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94138245; PubMed=8305874;  
RA Li J., Goldschmidt-Clermont M., Timko M.;  
RT "Chloroplast-encoded chlB is required for light-independent  
RT protochlorophyllide reductase activity in *Chlamydomonas reinhardtii*.";  
RL Plant Cell 5:1817-1829(1993).  
DR EMBL; D17510; BAA04307.1; -.  
DR EMBL; D11467; BAA02023.1; -.  
KW Chloroplast.  
SQ SEQUENCE 22 AA; 2606 MW; 0A95CB0443BCAEC5 CRC64;



Q37112 Length: 22 December 22, 2002 19:30 Type: P Check: 9119 ..

1 MEYLTTDRSI ECGIYLLKKIE SI

!!AA\_SEQUENCE 1.0  
ID Q957T4 PRELIMINARY; PRT; 22 AA.  
AC Q957T4;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Ribosomal protein S12 (Fragment).  
GN RPS12.  
OS Abies alba (Edeltanne) (European silver fir).  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
OX NCBI\_TaxID=45372;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Liepelt S.;  
RT "Sequence analysis of chosen regions of the mitochondrial genome of  
Abies alba.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF359456; AAK48935.1; -.  
KW Mitochondrion.  
FT NON\_TER 22 22  
SQ SEQUENCE 22 AA; 2711 MW; 97BF270D8C4E6558 CRC64;

Q957T4 Length: 22 December 22, 2002 19:30 Type: P Check: 9875 ..

1 MPTSNQSIRH GREKKRRTDR TR

!!AA\_SEQUENCE 1.0  
ID Q957T0 PRELIMINARY; PRT; 23 AA.  
AC Q957T0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Ribosomal protein S12 (Fragment).  
GN RPS12.  
OS Pinus mugo.  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_TaxID=28528;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wachowiak W., Liepelt S., Prus-Glowacki W.;  
RT "Sequence analysis of chosen regions of the mitochondrial genome of  
Pinus mugo and P. sylvestris.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF359576; AAK39117.1; -.  
KW Mitochondrion.  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270D8C4E65 CRC64;

Q957T0 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSIRH GREKKRRTDR TRA

!!AA\_SEQUENCE 1.0  
ID Q957S8 PRELIMINARY; PRT; 23 AA.  
AC Q957S8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Ribosomal protein S12 (Fragment).  
GN RPS12.  
OS Pinus sylvestris (Scots pine).  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_TaxID=3349;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wachowiak W., Liepelt S., Prus-Glowacki W.;  
RT "Sequence analysis of chosen regions of the mitochondrial genome of  
RT Pinus mugo and P. sylvestris.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF359577; AAK39119.1; -.  
KW Mitochondrion.  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2782 MW; 17A7BF270D8C4E65 CRC64;

Q957S8 Length: 23 December 22, 2002 19:30 Type: P Check: 1370 ..

1 MPTSNQSIRH GREKKRRTDR TRA

!!AA\_SEQUENCE 1.0  
ID Q9T2K6 PRELIMINARY; PRT; 20 AA.  
AC Q9T2K6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE Acyl-(acyl-carrier-protein) hydrolase 33 kDa polypeptide, AH1  
(EC 3.1.2.14) (Fragment).  
OS Cucurbita moschata (Cushaw squash) (Winter crookneck squash).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.  
OX NCBI\_TaxID=3662;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93004473; PubMed=1391766;  
RA Imai H., Nishida I., Murata N.;  
RT "Acyl-(acyl-carrier protein) hydrolase from squash cotyledons specific  
RT to long-chain fatty acids: purification and characterization.";  
RL Plant Mol. Biol. 20:199-206(1992).  
SQ SEQUENCE 20 AA; 1973 MW; AE5304F9CD73D36E CRC64;

Q9T2K6 Length: 20 December 22, 2002 19:30 Type: P Check: 5922 ..

1 GSSSLADRLX LGSLAXDGFS

!!AA\_SEQUENCE 1.0  
ID Q9T2H3 PRELIMINARY; PRT; 24 AA.  
AC Q9T2H3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Chaperonin 21 (Fragment).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95394976; PubMed=7665625;  
RA Ryan M.T., Naylor D.J., Hoogenraad N.J., Hoj P.B.;  
RL J. Biol. Chem. 270:22037-22043(1995).  
SQ SEQUENCE 24 AA; 2642 MW; B797841E1005A51A CRC64;

Q9T2H3 Length: 24 December 22, 2002 19:30 Type: P Check: 2853 ..

1 ATVVAPKYTS IKPTADRVL1 KIKE

!!AA\_SEQUENCE 1.0  
ID Q37852 PRELIMINARY; PRT; 18 AA.  
AC Q37852;



DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE A protein (Fragment).  
OS Bacteriophage R17.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;  
OC Levivirus.  
OX NCBI\_TaxID=12026;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=74125852; PubMed=4361645;  
RA Rensing U.F.E., Coulson A., Schoenmakers J.G.G.;  
RT "A sequence of 54 nucleotides from the A-protein cistron of Coliphage-  
RT R17 RNA.";  
RL Eur. J. Biochem. 41:431-438 (1974).  
DR EMBL; M24814; AAA32177.1; -.  
FT NON\_TERMINAL 1 1  
FT NON\_TERMINAL 18 18  
SQ SEQUENCE 18 AA; 2217 MW; CE3DE9FD8852F11E CRC64;

Q37852 Length: 18 December 22, 2002 19:30 Type: P Check: 3123 ..

1 ALRYLALNED RKFRSKHV

!!AA\_SEQUENCE 1.0  
ID Q42209 PRELIMINARY; PRT; 28 AA.  
AC Q42209;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Ubiquitin conjugating enzyme UBC10 (Fragment).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA; TISSUE=SEEDLING;  
RA Desprez T., Amselem J., Chiapello H., Caboche M., Hofte H.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; Z29145; CAA82396.1; -.  
DR HSSP; P15731; 1QCQ.  
DR InterPro; IPR000608; UBQ\_conjugat.  
DR Pfam; PF00179; UQ\_con; 1.  
DR ProDom; PD000461; UBQ\_conjugat; 1.  
FT NON\_TERMINAL 1 1  
SQ SEQUENCE 28 AA; 3321 MW; C4000F6BBD329BAF CRC64;

Q42209 Length: 28 December 22, 2002 19:30 Type: P Check: 1580 ..

1 PEIAHMYKTD RAKYESTARS WTQKYAMG

!!AA\_SEQUENCE 1.0  
ID Q9FZP4 PRELIMINARY; PRT; 29 AA.  
AC Q9FZP4;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Genomic DNA, chromosome 5, P1 clone:MVP2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";



RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AB025636; BAB11485.1; -.

SQ SEQUENCE 29 AA; 3557 MW; 9B75DC6B293119A2 CRC64;

Q9FZP4 Length: 29 December 22, 2002 19:30 Type: P Check: 3486 ..

1 MDDVDDVFSY LLSKEIDEEN EDREPKYVY

!!AA\_SEQUENCE 1.0

ID Q39633 PRELIMINARY; PRT; 23 AA.

AC Q39633;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)

DE Catalase (Fragment).

OS Cucumis sativus (Cucumber).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.

OX NCBI\_TaxID=3659;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=COTYLEDON;

RX MEDLINE=96104306; PubMed=8564304;

RA Toyama T., Teramoto H., Takeba G., Tsuji H.;

RT "Cytokinin induces a rapid decrease in the levels of mRNAs for

RT catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and other

RT unidentified proteins in etiolated cotyledons of cucumber.";

RL Plant Cell Physiol. 36:1349-1359(1995).

DR EMBL; D63385; BAA09701.1; -.

FT NON\_TER 1 1

SQ SEQUENCE 23 AA; 2445 MW; DD39DAFD58C3AE65 CRC64;

Q39633 Length: 23 December 22, 2002 19:30 Type: P Check: 1444 ..

1 NGSQADRSVG QKLAPHLNVR PSI

!!AA\_SEQUENCE 1.0

ID P82195 PRELIMINARY; PRT; 26 AA.

AC P82195;

DT 01-JUN-2000 (TrEMBLrel. 14, Created)

DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE Chloroplast 50S ribosomal protein L18 beta (Fragment).

OS Spinacia oleracea (Spinach).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.

OX NCBI\_TaxID=3562;

RN [1]

RP SEQUENCE.

RC STRAIN=CV. ALWARO; TISSUE=LEAF;

RX MEDLINE=20435798; PubMed=10874046;

RA Yamaguchi K., Subramanian A.R.;

RT "The plastid ribosomal proteins. Identification of all the proteins in  
RT the 50 S subunit of an organelle ribosome (chloroplast).";

RL J. Biol. Chem. 275:28466-28482(2000).

CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 5S RIBOSOMAL RNA.

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.

CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 KDA.

CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.

KW Ribosomal protein; Chloroplast; rRNA-binding.

FT NON\_TER 26 26

SQ SEQUENCE 26 AA; 3184 MW; C5664EC1B2AB7C97 CRC64;

P82195 Length: 26 December 22, 2002 19:30 Type: P Check: 7110 ..

1 KAHTRREDRT ARHVRIRKKV EGTPER

!!AA\_SEQUENCE 1.0



ID P82196 PRELIMINARY; PRT; 29 AA.  
AC P82196;  
DT 01-JUN-2000 (TrEMBLrel. 14, Created)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Chloroplast 50S ribosomal protein L18 alpha (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. ALWARO; TISSUE=LEAF;  
RX MEDLINE=20435798; PubMed=10874046;  
RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28482(2000).  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 5S RIBOSOMAL RNA.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 11.5 KDA.  
CC -!- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
FT NON\_TER 29 29  
SQ SEQUENCE 29 AA; 3517 MW; 86207D55664EC1B2 CRC64;

P82196 Length: 29 December 22, 2002 19:30 Type: P Check: 4154 ..

1 KAHTRREDRT ARHVRIRKKV EGTPERXXL

!!AA\_SEQUENCE 1.0  
ID P83089 PRELIMINARY; PRT; 19 AA.  
AC P83089;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Thylakoid luminal 25.3 kDa protein (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE, SUBCELLULAR LOCATION, AND MASS SPECTROMETRY.  
RC TISSUE=LEAF;  
RX PubMed=11719511;  
RA Schubert M., Petersson U.A., Haas B.J., Funk C., Schroeder W.P.,  
RA Kieselbach T.;  
RT "Proteome map of the chloroplast lumen of Arabidopsis thaliana.";  
RL J. Biol. Chem. 277:8354-8365(2002).  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN.  
CC -!- MASS SPECTROMETRY: MW=25300; METHOD=MALDI.  
KW Chloroplast; Thylakoid.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2093 MW; EC33422F8633F1F9 CRC64;

P83089 Length: 19 December 22, 2002 19:30 Type: P Check: 4307 ..

1 AIANAPLLDT TITDRVFFD

!!AA\_SEQUENCE 1.0  
ID Q9S8D2 PRELIMINARY; PRT; 29 AA.  
AC Q9S8D2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CMETI-B=TRYPSIN inhibitor.  
OS Cucumis melo (Muskmelon).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;



OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
OX NCBI\_TaxID=3656;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96015144; PubMed=8537309;  
RA Lee C.F., Lin J.Y.;  
RT "Amino acid sequences of trypsin inhibitors from the melon Cucumis  
RT melo.";  
RL J. Biochem. 118:18-22(1995).  
DR HSSP; P12071; 2ETI.  
DR InterPro; IPR000737; Squash.  
DR Pfam; PF00299; squash; 1.  
DR PRINTS; PR00293; SQUASHINHBTR.  
DR ProDom; PD003401; Squash; 1.  
DR SMART; SM00286; PTI; 1.  
DR PROSITE; PS00286; SQUASH\_INHIBITOR; 1.  
SQ SEQUENCE 29 AA; 3196 MW; A4BCFF7AA1AC300E CRC64;

Q9S8D2 Length: 29 December 22, 2002 19:30 Type: P Check: 2605 ..

1 VGCPRILMKC KTDRDCLTGC TCKRNGYCG

!!AA\_SEQUENCE 1.0  
ID Q9S898 PRELIMINARY; PRT; 16 AA.  
AC Q9S898;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE 23 kDa heat-induced protein (Fragment).  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96351184; PubMed=8742333;  
RA Sabehat A., Weiss D., Lurie S.;  
RT "The correlation between heat-shock protein accumulation and  
RT persistence and chilling tolerance in tomato fruit.";  
RL Plant Physiol. 110:531-537(1996).  
SQ SEQUENCE 16 AA; 1840 MW; 9B0D3F77BFEE6CE3 CRC64;

Q9S898 Length: 16 December 22, 2002 19:30 Type: P Check: 484 ..

1 NTNTQMTAYD QDDRGX

!!AA\_SEQUENCE 1.0  
ID Q9QVF2 PRELIMINARY; PRT; 12 AA.  
AC Q9QVF2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE TRANSFERRIN=PEPTIDE 21 (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92165927; PubMed=1791188;  
RA Cavanaugh P.G., Nicolson G.L.;  
RT "Lung-derived growth factor that stimulates the growth of lung-  
RT metastasizing tumor cells: identification as transferrin.";  
RL J. Cell. Biochem. 47:261-271(1991).  
FT NON\_TER 1 1  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1436 MW; 40AD1DFA420AADD3 CRC64;

Q9QVF2 Length: 12 December 22, 2002 19:30 Type: P Check: 5863 ..



1 NLANKADRDQ YE

!!AA\_SEQUENCE 1.0  
ID Q9CTN0 PRELIMINARY; PRT; 18 AA.  
AC Q9CTN0;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE B230218P12Rik protein (Fragment).  
GN B230218P12RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=CORPORA QUADRIGEMINA;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK021007; BAB32275.1; -.  
DR MGD; MGI:1926130; B230218P12Rik.  
FT NON\_TER 18 18  
SQ SEQUENCE 18 AA; 2136 MW; 72A0BE772BCA48C8 CRC64;

Q9CTN0 Length: 18 December 22, 2002 19:30 Type: P Check: 2989 ..

1 MPVKKKDTDR ALSLLEEY

!!AA\_SEQUENCE 1.0  
ID Q9CS32 PRELIMINARY; PRT; 27 AA.  
AC Q9CS32;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE 2610020J05Rik protein (Fragment).  
GN 2610020J05RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,



RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK019213; BAB31604.1; -.  
DR MGD; MGI:1914299; 2610020J05Rik.  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3299 MW; 95C1A07254A2C489 CRC64;

Q9CS32 Length: 27 December 22, 2002 19:30 Type: P Check: 8838 ..

1 MAENTDRNQI EKLLNRVKEL EQEVERL

!!AA\_SEQUENCE 1.0  
ID Q63990 PRELIMINARY; PRT; 25 AA.  
AC Q63990;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hox3.5 homeobox homolog protein (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WISTAR;  
RC TISSUE=BONE MORPHOGENETIC PROTEIN-IMPLANTED SUBCUTANEOUS MUSCLE;  
RX MEDLINE=94271262; PubMed=7911662;  
RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;  
RT "Changes in homeobox-containing gene expression during ectopic bone  
formation induced by bone morphogenetic protein.";  
RL Biochem. Biophys. Res. Commun. 201:980-987(1994).  
DR EMBL; S71286; AAB31006.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR ProDom; PD000010; Homeobox; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 3111 MW; 008AA81607BD2203 CRC64;

Q63990 Length: 25 December 22, 2002 19:30 Type: P Check: 5506 ..

1 LFNMYLTRER RLEISKTINL TDRQV

!!AA\_SEQUENCE 1.0  
ID O88226 PRELIMINARY; PRT; 33 AA.  
AC O88226;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Mszf76 (Fragment).  
GN MSZF76.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL6;  
RX MEDLINE=98296253; PubMed=9630514;  
RA Agata Y., Matsuda E., Shimizu A.;



RT "Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc  
RT finger proteins by degenerate PCR.";  
RL Gene 213:55-64(1998).  
DR EMBL; AB010321; BAA31377.1; -.  
DR InterPro; IPR000822; Znf\_C2H2.  
DR Pfam; PF00096; zf-C2H2; 1.  
DR ProDom; PD000003; Znf\_C2H2; 1.  
DR SMART; SM00355; ZnF\_C2H2; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 1.  
KW DNA-binding; Metal-binding; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 33 33  
SQ SEQUENCE 33 AA; 3730 MW; D8D31E5B5C54C538 CRC64;

088226 Length: 33 December 22, 2002 19:30 Type: P Check: 2429 ..

1 IHTGEKPYRC AECGKAFTDR SNLIKHQTH TGE

!!AA\_SEQUENCE 1.0  
ID Q9JIU0 PRELIMINARY; PRT; 32 AA.  
AC Q9JIU0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Protocadherin-T5 (Fragment).  
GN PCDH-T5.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FISCHER344; TISSUE=TESTIS;  
RX MEDLINE=20114371; PubMed=10650949;  
RA Johnson K.J., Patel S.R., Boekelheide K.;  
RT "Multiple cadherin superfamily members with unique expression profiles  
RT are produced in rat testis.";  
RL Endocrinology 141:675-683(2000).  
DR EMBL; AF177699; AAF87074.1; -.  
DR InterPro; IPR002126; Cadherin.  
DR Pfam; PF00028; cadherin; 1.  
FT NON\_TER 1 1  
FT NON\_TER 32 32  
SQ SEQUENCE 32 AA; 3369 MW; A6159FED42844051 CRC64;

✓

Q9JIU0 Length: 32 December 22, 2002 19:30 Type: P Check: 1050 ..

1 ALQAFEFHVVG ATDRGSPALS SQALVRVVVL DN

!!AA\_SEQUENCE 1.0  
ID Q9QX46 PRELIMINARY; PRT; 25 AA.  
AC Q9QX46;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Amyloid protein (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SVJ; TISSUE=LIVER;  
RX MEDLINE=20102720; PubMed=10636908;  
RA Bergsdorf C., Paliga K., Kreger S., Masters C.L., Beyreuther K.;  
RT "Identification of cis-Elements Regulating Exon 15 Splicing of the  
RT Amyloid Precursor Protein Pre-mRNA.";  
RL J. Biol. Chem. 275:2046-2056(2000).  
DR EMBL; AF199005; AAF20194.1; -.  
DR EMBL; AF199003; AAF20194.1; JOINED.

✓



DR EMBL; AF199004; AAF20194.1; JOINED.  
FT NON\_TER 1 1  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 2608 MW; 276F026D582421D0 CRC64;

Q9QX46 Length: 25 December 22, 2002 19:30 Type: P Check: 4770 ..

1 ENEVEPVNDAR PAADRGLTTR PGSQL

!!AA\_SEQUENCE 1.0  
ID Q9ET00 PRELIMINARY; PRT; 20 AA.  
AC Q9ET00;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE EIF4H (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RA Green E.D.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF289665; AAF99335.1; -.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2188 MW; 8AFFAC0FB953E8F5 CRC64;

Q9ET00 Length: 20 December 22, 2002 19:30 Type: P Check: 5807 ..

1 MADFDTYDDR AYSSFGGGRG

!!AA\_SEQUENCE 1.0  
ID Q9Z1I5 PRELIMINARY; PRT; 19 AA.  
AC Q9Z1I5;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE Calcineurin A alpha (EC 3.1.3.16) (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=93111954; PubMed=1335233;  
RA Chang C., Takeda T., Mukai H., Shuntoh H., Kuno T., Tanaka C.;  
RT "Molecular cloning and characterization of the promoter region of the  
calcineurin A alpha gene.";  
RL Biochem. J. 288:801-805(1992).  
DR EMBL; D10480; BAA01283.1; -.  
KW Hydrolase.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2115 MW; F905F8B134CCEC57 CRC64;

Q9Z1I5 Length: 19 December 22, 2002 19:30 Type: P Check: 4927 ..

1 MSEPKAIDPK LSTTDRVVK

!!AA\_SEQUENCE 1.0  
ID Q62256 PRELIMINARY; PRT; 18 AA.  
AC Q62256;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Spermatogenic-specific proenkephalin.  
GN PENK-RS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90287163; PubMed=2355920;  
RA Kilpatrick D.L., Zinn S.A., Fitzgerald M., Higuchi H., Sabol S.L.,  
RA Meyerhardt J.;  
RT "Transcription of the rat and mouse proenkephalin genes is initiated  
RT at distinct sites in spermatogenic and somatic cells.";  
RL Mol. Cell. Biol. 10:3717-3726(1990).  
DR EMBL; M55181; AAA40127.1; -.  
DR MGD; MGI:104628; Penk-rs.  
SQ SEQUENCE 18 AA; 2043 MW; B96E10CC7049FA76 CRC64;

Q62256 Length: 18 December 22, 2002 19:30 Type: P Check: 3193 ..

1 MSSGKQDSPW EDRIPPGR

!!AA\_SEQUENCE 1.0  
ID Q63985 PRELIMINARY; PRT; 25 AA.  
AC Q63985;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hox1.8 homeobox homolog protein (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94271262; PubMed=7911662;  
RA Iimura T., Oida S., Takeda K., Maruoka Y., Sasaki S.;  
RT "Changes in homeobox-containing gene expression during ectopic bone  
RT formation induced by bone morphogenetic protein.";  
RL Biochem. Biophys. Res. Commun. 201:980-987(1994).  
DR EMBL; S71275; AAB31001.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR ProDom; PD000010; Homeobox; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 25 AA; 3134 MW; C54E951607BD27B1 CRC64;

Q63985 Length: 25 December 22, 2002 19:30 Type: P Check: 5721 ..

1 LFNMYLTRER RLEISRSVHL TDRQV

!!AA\_SEQUENCE 1.0  
ID Q61159 PRELIMINARY; PRT; 32 AA.  
AC Q61159;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE NN8-AG (Fragment).  
GN RRG1 OR NN8-AG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SV;  
RX MEDLINE=96315660; PubMed=8754834;  
RA Shago M., Giguere V.;  
RT "Isolation of a novel retinoic acid-responsive gene by selection of  
RT genomic fragments derived from CpG-island-enriched DNA.";  
RL Mol. Cell. Biol. 16:4337-4348(1996).  
DR EMBL; U50384; AAB38132.1; -.  
DR MGD; MGI:108048; Rrg1.  
FT NON\_TER 32 32



SQ SEQUENCE 32 AA; 3470 MW; 9571FBD80B1FCA84 CRC64;

Q61159 Length: 32 December 22, 2002 19:30 Type: P Check: 837 ..

1 MAASMCDVFS FCGVGADRAR GSVEVRYVDS IK

!!AA\_SEQUENCE 1.0

ID Q923H1 PRELIMINARY; PRT; 40 AA.

AC Q923H1;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Aristaless-related homeobox protein Arx (Fragment).

OS Cricetulus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OC Cricetulus.

OX NCBI\_TaxID=10029;

RN [1]

RP SEQUENCE FROM N.A.

RA Ohira R.H., Zhang Y.H., Guo W., Dipple K., Shih S., Doerr J.,

RA Huang B.-L., Fu L., Abu-Khalil A., Geschwind D., McCabe E.;

RT "Human ARX gene: genomic characterization and expression.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AY038070; AAK93900.1; -.

KW DNA-binding; Homeobox; Nuclear protein.

FT NON\_TER 1 1

FT NON\_TER 40 40

SQ SEQUENCE 40 AA; 3815 MW; 87A852768DB0CB6E CRC64;

Q923H1 Length: 40 December 22, 2002 19:30 Type: P Check: 2032 ..

1 PTPAVEGAVA SGALADPATA AADRRASSIA AXGSRPRSMP

!!AA\_SEQUENCE 1.0

ID Q99KX5 PRELIMINARY; PRT; 27 AA.

AC Q99KX5;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Hypothetical 2.8 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; BC003968; AAH03968.1; -.

KW Hypothetical protein.

SQ SEQUENCE 27 AA; 2832 MW; 4008F368AF868B63 CRC64;

Q99KX5 Length: 27 December 22, 2002 19:30 Type: P Check: 9106 ..

1 MASLPVVGSQ VPADRGYLG PPAVQEH

!!AA\_SEQUENCE 1.0

ID Q9QVD0 PRELIMINARY; PRT; 24 AA.

AC Q9QVD0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Vitronectin (Fragment).

OS Cavia (guinea pigs).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.

OX NCBI\_TaxID=10140;

RN [1]

RP SEQUENCE.

RX MEDLINE=92207982; PubMed=1372829;



RA Nakashima N., Miyazaki K., Ishikawa M., Yatohgo T., Ogawa H.,  
RA Uchibori H., Matsumoto I., Seno N., Hayashi M.;  
RT "Vitronectin diversity in evolution but uniformity in ligand binding  
RT and size of the core polypeptide.";  
RL Biochim. Biophys. Acta 1120:1-10(1992).  
FT NON\_TER 1 1  
FT NON\_TER 24 24  
SQ SEQUENCE 24 AA; 2736 MW; 0CD91505B6F22D0A CRC64;

Q9QVD0 Length: 24 December 22, 2002 19:30 Type: P Check: 3385 ..

1 XXESXKGRXT EGFNADRDXQ XXEL

!!AA\_SEQUENCE 1.0  
ID Q9QV89 PRELIMINARY; PRT; 32 AA.  
AC Q9QV89;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE SLP-14=FATTY acid-binding protein (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93191717; PubMed=8447836;  
RA Bansal M.P., Medina D.;  
RT "Expression of fatty acid-binding proteins in the developing mouse  
RT mammary gland.";  
RL Biochem. Biophys. Res. Commun. 191:61-69(1993).  
DR HSSP; P05413; 1HMT.  
DR InterPro; IPR000566; Lipocln\_cytFABP.  
DR Pfam; PF00061; lipocalin; 1.  
FT NON\_TER 1 1  
FT NON\_TER 32 32  
SQ SEQUENCE 32 AA; 3495 MW; 67D23A0146E46AFE CRC64;

Q9QV89 Length: 32 December 22, 2002 19:30 Type: P Check: 581 ..

1 EISFQLGVEF DEVTADDRKV KSVVTLGGK LV

!!AA\_SEQUENCE 1.0  
ID Q9QV71 PRELIMINARY; PRT; 17 AA.  
AC Q9QV71;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE Lactase-phlorizin hydrolase (Fragment).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93293888; PubMed=8514793;  
RA Dudley M.A., Hachey D.L., Quaroni A., Hutchens T.W., Nichols B.L.,  
RA Rosenberger J., Perkinson J.S., Cook G., Reeds P.J.;  
RT "In vivo sucrase-isomaltase and lactase-phlorizin hydrolase turnover  
RT in the fed adult rat.";  
RL J. Biol. Chem. 268:13609-13616(1993).  
SQ SEQUENCE 17 AA; 1947 MW; 08941FD316F9692F CRC64;

Q9QV71 Length: 17 December 22, 2002 19:30 Type: P Check: 1363 ..

1 DWEDRNFIAA GPLTNDL

!!AA\_SEQUENCE 1.0  
ID Q61461 PRELIMINARY; PRT; 25 AA.  
AC Q61461;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)



DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Cytochrome P-450b (Phenobarbital-inducible) (Fragment).  
GN CYP2B13.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84207435; PubMed=6547088;  
RA Stupans I., Ikeda T., Kessler D.J., Nebert D.W.;  
RT "Characterization of a cDNA clone for mouse phenobarbital-inducible  
RT cytochrome P-450b.";  
RL DNA 3:129-137(1984).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; K02409; AAA37510.1; -.  
DR MGD; MGI:88599; Cyp2b13.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; p450; 1.  
KW Heme; Monooxygenase; Oxidoreductase.  
FT NON\_TER 1 1  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 2979 MW; B9DFF7A57355A1B0 CRC64;

Q61461 Length: 25 December 22, 2002 19:30 Type: P Check: 4925 ..

1 SHRLPTLDDR SKMPYTDavi HEIQR

!!AA\_SEQUENCE 1.0  
ID Q65747 PRELIMINARY; PRT; 19 AA.  
AC Q65747;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Outer coat protein (VP2) (Fragment).  
OS Bluetongue virus.  
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.  
OX NCBI\_TaxID=12591;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88265863; PubMed=2838961;  
RA Gould A.R., Hyatt A.D., Eaton B.T.;  
RT "Morphogenesis of a bluetongue virus variant with an amino acid  
RT alteration at a neutralization site in the outer coat protein, VP2.";  
RL Virology 165:23-32(1988).  
DR EMBL; M21355; AAA42845.1; -.  
DR InterPro; IPR001742; Orbi\_VP2.  
DR Pfam; PF00898; Orbi\_VP2; 1.  
KW Coat protein.  
FT NON\_TER 1 1  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2608 MW; F0428841265D3061 CRC64;

Q65747 Length: 19 December 22, 2002 19:30 Type: P Check: 4484 ..

1 ERLKIFEHRN QRRDEDRFY

!!AA\_SEQUENCE 1.0  
ID Q66538 PRELIMINARY; PRT; 35 AA.  
AC Q66538;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE 3' proximal protein (Fragment).  
OS Ebola virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;  
OC Ebola-like viruses.  
OX NCBI\_TaxID=11268;  
RN [1]  
RP SEQUENCE FROM N.A.



RX MEDLINE=86124724; PubMed=3946083;  
RA Kiley M.P., Wilusz J., McCormick J.B., Keene J.D.;  
RT "Conservation of the 3' terminal nucleotide sequences of Ebola and  
Marburg virus.";  
RL Virology 149:251-254(1986).  
DR EMBL; M33062; AAA42976.1; -.  
FT NON\_TER 35 35  
SQ SEQUENCE 35 AA; 4142 MW; F7778E5FC6283092 CRC64;

Q66538 Length: 35 December 22, 2002 19:30 Type: P Check: 8137 ..

1 MRKINNFLSL KFDDRNLKLK LLICNHTVDS EPHTS

!!AA\_SEQUENCE 1.0  
ID Q83622 PRELIMINARY; PRT; 9 AA.  
AC Q83622;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 1.1 kDa protein (Fragment).  
OS Murray valley encephalitis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Flavivirus.  
OX NCBI\_TaxID=11079;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88118912; PubMed=2828633;  
RA Hahn C.S., Hahn Y.S., Rice C.M., Lee E., Dalgarno L., Strauss E.G.,  
RA Strauss J.H.;  
RT "Conserved elements in the 3' untranslated region of flavivirus RNAs  
RT and potential cyclization sequences.";  
RL J. Mol. Biol. 198:33-41(1987).  
DR EMBL; M35172; AAA66627.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 1055 MW; FF36D40AAB05A2C1 CRC64;

Q83622 Length: 9 December 22, 2002 19:30 Type: P Check: 3517 ..

1 THVSEDRVL

!!AA\_SEQUENCE 1.0  
ID Q84254 PRELIMINARY; PRT; 22 AA.  
AC Q84254;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE X protein (Fragment).  
OS Bovine papillomavirus.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10571;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89067912; PubMed=2848926;  
RA Stamps A.C., Campo M.S.;  
RT "Mapping of two novel transcripts of Bovine papillomavirus type 4.";  
RL J. Gen. Virol. 69:3033-3045(1988).  
DR EMBL; M35264; AAA46926.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 22 22  
SQ SEQUENCE 22 AA; 2273 MW; 2801BC23480C9CF9 CRC64;

Q84254 Length: 22 December 22, 2002 19:30 Type: P Check: 9066 ..

1 NAGPKPGTTP EDVADRPPDL PE

!!AA\_SEQUENCE 1.0  
ID Q9PWC1 PRELIMINARY; PRT; 39 AA.  
AC Q9PWC1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)



DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE P55-related MAGUK protein DLG3 (Fragment).  
GN DLG3.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Koenig C., Yan Y.L., Postlethwait J., Wendler S., Campos-Ortega J.A.;  
RT "A recessive mutation leading to vertebral ankylosis is associated  
RT with amino acid exchanges in the zebrafish homologue of the human  
RT membrane associated guanylate kinase protein DLG3.";  
RL Mech. Dev. 84:1-12(1999).  
DR EMBL; AF124436; AAD39393.1; -.  
DR InterPro; IPR004172; L27.  
DR Pfam; PF02828; L27; 1.  
FT NON\_TER 1 1  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4477 MW; 6A9D5C428F5A2DBF CRC64;

Q9PWC1 Length: 39 December 22, 2002 19:30 Type: P Check: 9645 ..

1 LYELLSVLPS QLQPHVESPD DRSFLHAMFG ERSLHSLVK

!!AA\_SEQUENCE 1.0  
ID 093438 PRELIMINARY; PRT; 29 AA.  
AC 093438;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Lbx1 protein (Fragment).  
GN LBX1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dietrich S., Schubert F.R., Healy C., Sharpe P.T., Lumsden A.;  
RT "Specification of the hypaxial musculature.";  
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Schubert F.R., Dietrich S., Chapman S.C., Lumsden A.;  
RT "Expression of the Lbx1 gene in the chick embryo.";  
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ222831; CAA11016.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 29 AA; 3352 MW; CABBC02415F1D4F7 CRC64;

093438 Length: 29 December 22, 2002 19:30 Type: P Check: 3493 ..

1 RFLYQKYLSP ADRDQIAQQL GLTNAQVIT

!!AA\_SEQUENCE 1.0  
ID Q90297 PRELIMINARY; PRT; 27 AA.  
AC Q90297;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Homeobox (Fragment).  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;



OC Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=RETINA;  
RX MEDLINE=93219357; PubMed=8096640;  
RA Levine E.M., Schechter N.;  
RT "Homeobox genes are expressed in the retina and brain of adult  
RT goldfish.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).  
DR EMBL; L09693; AAA49178.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEOBOX.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

Q90297 Length: 27 December 22, 2002 19:30 Type: P Check: 9410 ..

1 LFNMYLTRER RLEISKSINL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID Q90298 PRELIMINARY; PRT; 27 AA.  
AC Q90298;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Homeobox (Fragment).  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=RETINA;  
RX MEDLINE=93219357; PubMed=8096640;  
RA Levine E.M., Schechter N.;  
RT "Homeobox genes are expressed in the retina and brain of adult  
RT goldfish.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).  
DR EMBL; L09694; AAA49179.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEOBOX.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3375 MW; 52B7254E951607BD CRC64;

Q90298 Length: 27 December 22, 2002 19:30 Type: P Check: 9642 ..

1 LFNMYLTRER RLEISRSVHL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID Q07145 PRELIMINARY; PRT; 27 AA.  
AC Q07145;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Homeodomain (Fragment).  
GN HOMEOBOX.  
OS Petromyzon marinus (Sea lamprey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
OX NCBI\_TaxID=7757;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93317669; PubMed=8101001;



RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;  
RT "Expansion of the Hox gene family and the evolution of chordates.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).  
DR EMBL; L14895; AAA02530.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3322 MW; 5529340A951607BD CRC64;

Q07145 Length: 27 December 22, 2002 19:30 Type: P Check: 9552 ..

1 LFNMYLTRER RLEISRGVNL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID Q07150 PRELIMINARY; PRT; 27 AA.  
AC Q07150;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Homeodomain (Fragment).  
GN HOMEobox.  
OS Petromyzon marinus (Sea lamprey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
OX NCBI\_TaxID=7757;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93317669; PubMed=8101001;  
RA Pendleton J.W., Nagai B.K., Murtha M.T., Ruddle F.H.;  
RT "Expansion of the Hox gene family and the evolution of chordates.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6300-6304(1993).  
DR EMBL; L14900; AAA02535.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3319 MW; 53CD02F1671607BD CRC64;

Q07150 Length: 27 December 22, 2002 19:30 Type: P Check: 9407 ..

1 LFSMYLTRER RLEISHLLSL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID O73591 PRELIMINARY; PRT; 14 AA.  
AC O73591;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hox C10 (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WHITE LEGHORN;  
RX MEDLINE=98141813; PubMed=9473273;  
RA Peale F.V. Jr., Mason K., Hunter A.W., Bothwell M.;  
RT "Multiplex display polymerase chain reaction amplifies and resolves  
RT related sequences sharing a single moderately conserved domain.";  
RL Anal. Biochem. 256:158-168(1998).  
DR EMBL; U34614; AAC36452.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 14 14  
SQ SEQUENCE 14 AA; 1673 MW; 81258FC9E81FDA45 CRC64;



073591 Length: 14 December 22, 2002 19:30 Type: P Check: 8165 ..

1 RLEISKSINL TDRQ

!!AA\_SEQUENCE 1.0  
ID P83009 PRELIMINARY; PRT; 13 AA.  
AC P83009;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Phospholemmann (FXYD domain-containing ion transport regulator 1)  
DE (Fragment).  
OS Lamna nasus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Lamniformes; Alopiidae;  
OC Lamna.  
OX NCBI\_TaxID=7849;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=RECTAL GLAND;  
RA Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.E.;  
RT "Phospholemmann in the rectal gland of sharks.";  
RL Submitted (JUN-2001) to the SWISS-PROT data bank.  
CC -!- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT  
CC WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN  
CC MUSCLE CONTRACTION.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSOMAL  
CC MEMBRANE.  
CC -!- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN  
CC KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT  
CC TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC  
CC STIMULATION (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE FXYD FAMILY.  
DR InterPro; IPRO00272; ATP1G1\_PLM\_MAT8.  
DR PROSITE; PS01310; FXYD; PARTIAL.  
KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;  
KW Microsome.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1542 MW; 081373C69724A444 CRC64;

P83009 Length: 13 December 22, 2002 19:30 Type: P Check: 7125 ..

1 VSDVPNNDDR FTY

!!AA\_SEQUENCE 1.0  
ID P83010 PRELIMINARY; PRT; 13 AA.  
AC P83010;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Phospholemmann (FXYD domain-containing ion transport regulator 1)  
DE (Fragment).  
OS Triakis scyllium (Leopard shark) (Triakis scyllia).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;  
OC Triakis.  
OX NCBI\_TaxID=30494;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=RECTAL GLAND;  
RA Schuurmans Stekhoven F.M.A.H., Flik G., Wendelaar Bonga S.E.;  
RT "Phospholemmann in the rectal gland of sharks.";  
RL Submitted (JUN-2001) to the SWISS-PROT data bank.  
CC -!- FUNCTION: INDUCES A HYPERPOLARIZATION-ACTIVATED CHLORIDE CURRENT  
CC WHEN EXPRESSED IN XENOPUS OOCYTES. MAY HAVE A FUNCTIONAL ROLE IN  
CC MUSCLE CONTRACTION.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MICROSOMAL  
CC MEMBRANE.  
CC -!- PTM: MAJOR PLASMA MEMBRANE SUBSTRATE FOR CAMP-DEPENDENT PROTEIN  
CC KINASE (PK-A) AND PROTEIN KINASE C (PK-C) IN SEVERAL DIFFERENT  
CC TISSUES. PHOSPHORYLATED IN RESPONSE TO INSULIN AND ADRENERGIC



CC STIMULATION (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE FXYD FAMILY.  
DR InterPro; IPR000272; ATP1G1\_PLM\_MAT8.  
DR PROSITE; PS01310; FXYD; PARTIAL.  
KW Transmembrane; Phosphorylation; Ionic channel; Ion transport;  
KW Microsome.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1497 MW; 1D140C92C1AE1444 CRC64;

P83010 Length: 13 December 22, 2002 19:30 Type: P Check: 6920 ..

1 AGEPANNEDR FNY

!!AA\_SEQUENCE 1.0  
ID Q90ZG3 PRELIMINARY; PRT; 27 AA.  
AC Q90ZG3;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE HoxC11a (Fragment).  
GN MF-HOXC11A.  
OS Oryzias latipes (Medaka fish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.  
OX NCBI\_TaxID=8090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kurosawa G., Yamada K., Ishiguro H., Hori H.;  
RT "isolation of BAC Clones That Cover Seven Hox Clusters in Medaka  
RT Genome.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB055740; BAB62853.1; -.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR ProDom; PD000010; Homeobox; 1.  
FT NON\_TER 1 1  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 3368 MW; 548C73EA983E9158 CRC64;

Q90ZG3 Length: 27 December 22, 2002 19:30 Type: P Check: 9516 ..

1 FFNVYINKEK RLQLSRMLNL TDRQVKI

!!AA\_SEQUENCE 1.0  
ID Q9PS38 PRELIMINARY; PRT; 20 AA.  
AC Q9PS38;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE PO GLYCOPROTEIN=MYELIN major structural protein (Fragment).  
OS Rana catesbeiana (Bull frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.  
OX NCBI\_TaxID=8400;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93002191; PubMed=1382532;  
RA Karthigasan J., Bauer T.K., Teplow D.B., Saavedra R.A.,  
RA Kirschner D.A.;  
RT "Phylogenetically conserved amino acids of MBP and PO from amphibian  
RT myelin.";  
RL J. Mol. Neurosci. 3:185-188(1992).  
FT NON\_TER 1 1  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2351 MW; AB238ACED7837676 CRC64;

Q9PS38 Length: 20 December 22, 2002 19:30 Type: P Check: 6634 ..

1 IEVYTDREIQ SNVGSKVHLY



!!AA\_SEQUENCE 1.0  
ID Q9PS32 PRELIMINARY; PRT; 23 AA.  
AC Q9PS32;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Nonmuscle alpha-actinin 115 kDa isoform (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93100311; PubMed=1334489;  
RA Imamura M., Masaki T.;  
RT "A novel nonmuscle alpha-actinin. Purification and characterization of  
chicken lung alpha-actinin.";  
RL J. Biol. Chem. 267:25927-25933(1992).  
FT NON\_TER 1 1  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2823 MW; EBED28437F668363 CRC64;

Q9PS32 Length: 23 December 22, 2002 19:30 Type: P Check: 1521 ..

1 LASDLLEWIR RTIPWLEDRS PQK

!!AA\_SEQUENCE 1.0  
ID Q9PRV3 PRELIMINARY; PRT; 24 AA.  
AC Q9PRV3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE NAP-185 NEUROGLIA-associated protein (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95050502; PubMed=7961672;  
RA Solca F.F., Lurie D.I., Diltz C.D., Johnson R.S., Kumar S.,  
RA Rubel E.W., Fischer E.H.;  
RT "Identification and purification of a chicken brain neuroglia-  
associated protein.";  
RL J. Biol. Chem. 269:27559-27565(1994).  
SQ SEQUENCE 24 AA; 2404 MW; 73E0E6067F253442 CRC64;

Q9PRV3 Length: 24 December 22, 2002 19:30 Type: P Check: 2489 ..

1 DGGEDRDAAV EEAVALGTGGC RTPK

!!AA\_SEQUENCE 1.0  
ID Q9DU42 PRELIMINARY; PRT; 25 AA.  
AC Q9DU42;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Rev protein (Fragment).  
GN REV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MIDU2;  
RX MEDLINE=21002573; PubMed=11118069;  
RA Motomura K., Kusagawa S., Kato K., Nohtomi K., Lwin H.H., Tun K.M.,  
RA Thwe M., Oo K.Y., Lwin S., Kyaw O., Zaw M., Nagai Y., Takebe Y.;  
RT "Emergence of new forms of human immunodeficiency virus type 1"



RT intersubtype recombinants in central myanmar.";  
RL AIDS Res. Hum. Retroviruses 16:1831-1843(2000).  
DR EMBL; AB043899; BAB19213.1; -.  
DR InterPro; IPR000625; REV\_protein.  
DR Pfam; PF00424; REV; 1.  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 2964 MW; 0F92F42A1DDF2758 CRC64;

Q9DU42 Length: 25 December 22, 2002 19:30 Type: P Check: 5339 ..

1 MAGRSEDRDE ELLKTVRLIK LLYQS

!!AA\_SEQUENCE 1.0  
ID Q70140 PRELIMINARY; PRT; 9 AA.  
AC Q70140;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Tat protein (Fragment).  
GN TAT.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=020;  
RX MEDLINE=95194694; PubMed=7888189;  
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,  
RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;  
RT "Genetic variation of HIV type 1 in four World Health Organization-  
RT sponsored vaccine evaluation sites: generation of functional envelope  
RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,  
RT and E. WHO Network for HIV Isolation and Characterization.";  
RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=020;  
RX MEDLINE=96190564; PubMed=8627686;  
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,  
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,  
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,  
RA Hahn B.H.;  
RT "Molecular cloning and analysis of functional envelope genes from  
RT human immunodeficiency virus type 1 sequence subtypes A through G. The  
RT WHO and NIAID Networks for HIV Isolation and Characterization.";  
RL J. Virol. 70:1651-1657(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=020;  
RA Allen E.E.;  
RL Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U08794; AAB05175.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 1098 MW; 5B76D40AB1AB01A3 CRC64;

Q70140 Length: 9 December 22, 2002 19:30 Type: P Check: 3335 ..

1 SKTETDRFD

!!AA\_SEQUENCE 1.0  
ID O25575 PRELIMINARY; PRT; 23 AA.  
AC O25575;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein HP0917.  
GN HP0917.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;



RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;'  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori.";  
RL Nature 388:539-547(1997).  
DR EMBL; AE000601; AAD07968.1; -.  
DR TIGR; HP0917; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 23 AA; 2717 MW; FC3042379DA2FD44 CRC64;

025575 Length: 23 December 22, 2002 19:30 Type: P Check: 575 ..

1 MSPLTPLRN P LTQEDRFFQE IIA

!!AA\_SEQUENCE 1.0  
ID 006283 PRELIMINARY; PRT; 27 AA.  
AC 006283;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein Rv3599c.  
GN RV3599C OR MTCTY07H7B.23.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
DR EMBL; Z95557; CAB08945.1; -.  
DR TubercuList; Rv3599c; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 27 AA; 2894 MW; 1746F2AA97B95C6B CRC64;

006283 Length: 27 December 22, 2002 19:30 Type: P Check: 8577 ..

1 MPASSLGTGS PAADRLDATH ERRREVI

!!AA\_SEQUENCE 1.0  
ID Q9KTZ5 PRELIMINARY; PRT; 33 AA.  
AC Q9KTZ5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical protein VC0735.  
GN VC0735.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.



OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae.";  
RL Nature 406:477-483(2000).  
DR EMBL; AE004159; AAF93900.1; -.  
DR TIGR; VC0735; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 33 AA; 3661 MW; A3AD77F43D6FE40C CRC64;

Q9KTZ5 Length: 33 December 22, 2002 19:30 Type: P Check: 3787 ..

1 MPVGSSIPCS ISTALADRKV VSRYWCFFTL LVL

!!AA\_SEQUENCE 1.0  
ID Q9KLI3 PRELIMINARY; PRT; 32 AA.  
AC Q9KLI3;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical protein VCA0761.  
GN VCA0761.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae.";  
RL Nature 406:477-483(2000).  
DR EMBL; AE004404; AAF96659.1; -.  
DR TIGR; VCA0761; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 32 AA; 3898 MW; E08F60B7C8C67F83 CRC64;

Q9KLI3 Length: 32 December 22, 2002 19:30 Type: P Check: 1300 ..

1 MPDRLLRFSA ICSTDRRKQT NLFSSDLKQQ WR

!!AA\_SEQUENCE 1.0  
ID Q8X407 PRELIMINARY; PRT; 13 AA.  
AC Q8X407;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein z0899.  
GN Z0899.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]



RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
DR EMBL; AE005252; AAG55068.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 13 AA; 1575 MW; 5127D365A6BBD763 CRC64;

Q8X407 Length: 13 December 22, 2002 19:30 Type: P Check: 6940 ..

1 MSTDRKPVML LFH

!!AA\_SEQUENCE 1.0  
ID Q8VJ09 PRELIMINARY; PRT; 35 AA.  
AC Q8VJ09;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein MT3536.  
GN MT3536.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / OSHKOSH;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AE007158; AAK47876.1; -.  
DR TIGR; MT3536; -.  
KW Hypothetical protein.  
SQ SEQUENCE 35 AA; 4204 MW; 897851C24908FC03 CRC64;

Q8VJ09 Length: 35 December 22, 2002 19:30 Type: P Check: 8356 ..

1 MKYKLAILDE YDRADRTERG AILRRENLYS SLLTE

!!AA\_SEQUENCE 1.0  
ID Q9HNL3 PRELIMINARY; PRT; 32 AA.  
AC Q9HNL3;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Vng2049c.  
GN VNG2049C.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,



RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL; AE005097; AAG20207.1; -.  
KW Complete proteome.  
SQ SEQUENCE 32 AA; 3719 MW; D0F9677E650A578A CRC64;

Q9HNL3 Length: 32 December 22, 2002 19:30 Type: P Check: 9872 ..

1 MVRVPVTDAG RAEMRERADR LETTAAFWRL VD

